Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

WILLIAM K. ESTES, Associate Editor
Indiana University

Consulting Editors

NORMAN H. ANDERSON, *University of California, Los Angeles*
E. JAMES ARCHER, *University of Wisconsin*
FRED ATTNEAVE, III, *University of Oregon*
CLETUS J. BURKE, *Indiana University*
JAMES DEESE, *Johns Hopkins University*
PAUL M. FITTS, *University of Michigan*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
HAROLD W. HAKE, *University of Illinois*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
HERSCHEL W. LEIBOWITZ, *International Business Machines Corporation, Bethesda, Maryland*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
LEO POSTMAN, *University of California, Berkeley*
L. STARLING REID, *University of Virginia*
KENNETH W. SPENCE, *State University of Iowa*
DELOS D. WICKENS, *Ohio State University*

CONTENTS

On the Inhibitory Effects of a Second Stimulus Following the Primary Stimulus to React: H. HELSON AND J. A. STEGER 201

Response Suppression in Perceptual Defense: R. B. ZAJONC 206

Factors in the Retention and Relearning of Perceptual-Motor Skill: E. A. FLEISHMAN AND J. F. PARKER, JR. 215

Discrimination of the Reward in Learning with Partial and Continuous Reinforcement: S. H. HULSE 227

Mediated Association in a Paired-Associate Transfer Task: D. S. PALERMO 234

Prediction of Preference, Transposition, and Transposition-Reversal from the Generalization Gradient: W. K. HONIG 239

Resistance to Extinction as a Function of Age and Schedules of Reinforcement: N. KASS 249

Subjective Scale of Force for a Large Muscle Group: H. EISLER 253

Re-examination of the Serial Position Effect: M. GLANZER AND S. C. PETERS 258

Effects of Instructions in Probability Learning: J. MCCRACKEN, C. OSTERHOUT, AND J. F. VOSS 267

Cognitive Factors in Heart Rate Conditioning: B. B. CHATTERJEE AND C. W. ERIKSEN 272

Test of the Hypothesis of Psychological Refractory Period: J. A. ADAMS 280

(Continued on inside back cover)

American Psychological Association

Vol. 64, No. 3 September 1962

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $10.00, or $20.00 annually. Single copies are $2.00. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This Journal publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research efforts is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this Journal, 1957, 53, 1–2).

Manuscripts *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc., as described in the *Publication Manual of the American Psychological Association.* When in doubt about practices of the Journal, authors should examine a recent issue.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1962 the publication cost is $30 per printed page.

All **regular** articles, **early publication** articles, **supplementary reports,** and **replication reports** should be submitted to the Editor-elect: David A. Grant: Department of Psychology, University of Wisconsin; Madison 6, Wisconsin. Manuscripts must be accompanied by an abstract of 100 to 120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts.* Detailed instructions for preparation of the abstracts may be obtained from the Editor or from the APA Central Office.

HELEN ORR
Managing Editor

ELIZABETH S. REED
Advertising Manager

RUTH P. JAGOLINZER
Editorial Assistant

Second-class postage paid at Lancaster, Pa.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the tenth of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1962 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 64, NO. 3 SEPTEMBER 1962

ON THE INHIBITORY EFFECTS OF A SECOND STIMULUS FOLLOWING THE PRIMARY STIMULUS TO REACT

HARRY HELSON AND JOSEPH A. STEGER

Kansas State University

Many conditions affecting simple spot reactions to visual, auditory, and tactile stimuli have been investigated (cf. Teichner, 1954) but, so far as we have been able to ascertain, the following phenomenon has not been hitherto reported in the literature: if a second stimulus *follows* the primary stimulus to react, reaction time (RT) is *lengthened* compared with RT when only a single stimulus is employed. Todd (1912) found that RT to a signal (visual, auditory, tactile) was *lengthened* when one or two other stimuli *preceded* the stimulus to react. The findings reported here also appear to be unique in that the intervals after onset of the primary stimulus (S_1) during which the second stimulus (S_2) still exerts a significant influence on RT are very long, considering the magnitude of RT to a single stimulus. In this study we shall be concerned with the case where S_1 and S_2 are both in the same sense modality (vision), leaving the report of an investigation of heteromodal effects to a later publication.[1]

[1] The phenomenon discussed in this article was discovered by the senior author in 1925 when attempting to condition sensory responses. In order to conceal the purpose of the earlier experiment from *Ss* they were required to react to a light (tone) which was followed by a tone (light) 75 msec. later in most of the trials. In some trials the second stimulus was omitted and it was then noticed that RTs were shorter when only a single stimulus was presented than when the second stimulus was given. We have now completed an investigation with light and tone as the stimuli with results similar to those reported here for unimodal stimulation.

METHOD

Subjects.—The *Ss* were 15 male students enrolled in an introductory psychology course ranging from 18 to 21 yr. of age. They were randomly divided into two groups, Group E (experimental), consisting of 10 *Ss*, and Group C (control), consisting of 5 *Ss*.

Apparatus and procedure.—The apparatus consisted of a black panel board on which were mounted three small neon lamps (GE 51) spaced ¾ in. apart horizontally with a small red fixation light slightly above the center lamp. A response button was conveniently placed for *S* to press as soon as the primary light (S_1) went on. The *E* sat behind the panel board and by pressing a single button actuated S_1, the Standard Electric clock which measured the reaction times in .01 sec., and the Hunter interval timer which operated the current on the second stimulus (S_2) at intervals ranging from 10 to 180 msec. in steps of 10 msec. Both S_1 and S_2 stayed on until *S* responded since it is known that duration of a flashing visual stimulus affects

RT. The room in which the experiment was conducted was dimly lighted and shielded from external sounds. The luminances of the stimuli were not determined but they were distinctly visible against the black panel in the dim light and were not noticeably different from one another.

The three visual stimuli were employed for the purpose of counterbalancing position of S_1 vis-à-vis S_2, the interpolated stimulus. S_1 was thus either on the right or the left of the middle stimulus, the latter always serving as S_2. The actual procedure can best be understood from the instructions given to *S*s:

> This is an experiment in simple reaction time. You are to respond to a light as quickly as you possibly can by pressing the button under your index finger.
>
> Note the lights facing you. I will tell you which end light you are to watch. If I say "left," you will watch the light on your left and react to it as quickly as possible. If I say "right," you will watch the end light to your right and react to it as quickly as possible.
>
> The procedure will be this: I will say "right" or "left" and then I will say "ready." A short time after I say "ready" the light you are to watch will come on.
>
> Are there any questions?

The right and left presentations of S_1 were randomly distributed and appeared from .5 to 3 sec. after the ready signal. The intervals at which S_2 followed S_1 were randomly sequenced. Each *S* in Group E reacted a total of 380 times, 20 times with S_2 presented at each of the 18 intervals and 20 times when S_2 was omitted, the latter being randomly interspersed among the trials in which S_2 followed the primary signal. The *S*s in Group C reacted 360 times to S_1 presented in random order on the right or on the left and with no S_2. Two control sets of RTs were thus available with which to compare the experimental RTs: (*a*) the RTs of Group E without S_2, and (*b*) the RTs of Group C with only S_1.

The *S*s were given four trials with the primary stimuli to familiarize them with the procedure. There was a 5-sec. break between trials and a 2-min. break at the end of Trial 180. The *S*s were not told the purpose of the experiment or the role of S_2. All parts of the equipment were shielded from *S*s except the light stimuli.

TABLE 1

TREND ANALYSIS OF REACTION TIME × INTERVALS DATA

Source of Variation	*df*	*MS*	*F*
Linear	1	26,332	2.183
Quadratic	1	74,077	6.142*
Cubic	1	2,533	.21
Intervals	17	44,653	3.703**
*S*s	9	1,292,452	
*S*s × I	153	12,060	
Residual	14	45,169	3.745**

* $P = .05$.
** $P = .01$.

RESULTS

The results leave no doubt as to the inhibiting effect of S_2 on RT to S_1. The results of a trend analysis (Grant, 1956) of the data across interstimulus intervals in Table 1 show that the quadratic component is significant at the .05 level and that the linear and cubic components are not significant.[2] This finding lends support to the expectation that the effect of S_2 should be minimal at some very short interval following S_1, that it should increase to some maximal value or values at certain interstimulus intervals, and then decline as the response is being consummated. Both this reasoning and the trend analysis suggest a parabola as the proper type of function to fit the data in Fig. 1. A second finding from the trend analysis is that the overall influence of interstimulus intervals is significant at the .01 level, which is confirmed by individual *t* tests of the differences between the mean RTs of Group E with and without S_2: 17 of the 18 interstimulus intervals yield significantly longer RTs at the following levels: 9 beyond .001, 4 beyond .01, 4 between .01 and .05, and 1 not significant (with a 180-msec. interval). The *t* tests were based on comparisons of means of 20

[2] We wish to express our gratitude to John Gaito for aid with the trend analysis and to John Gaito and to John Overall for discussion of a number of statistical issues that arose in connection with treatment of the data.

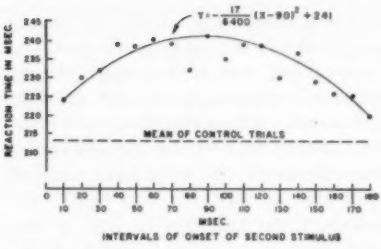


FIG. 1. Reaction time to a primary stimulus when a second stimulus follows at intervals ranging from 10 to 180 msec.

RTs by 10 *Ss* under the two experimental conditions using 9 *df* at each of the 18 interstimulus intervals. A second control, the mean RT of Group C which was given only S_1, was even lower than the mean RT in the control trials of Group E when S_2 was omitted ($t = 2.82$, $P < .01$), a finding that shows that there was some carry-over in Group E from the S_2 trials to the trials in which only S_1 was presented.

The curve in Fig. 1 represents a parabola made to fit the 10-, 90-, and 180-msec. intervals since it is at these times that the minimal and maximal inhibiting effects of S_2 are found under the conditions of this experiment. The RTs at 80, 100, and 130 msec. depart quite widely from the fitted curve and could undoubtedly have been approximated more closely if it had been fitted according to the least squares criterion which minimizes the total sum of squared differences between all observed and theoretical points. The continuous line in Fig. 1 is given by the equation:

$$Y = (-17/6400)(X - 90)^2 + 241$$

where Y is RT and X is the interval between S_1 and S_2. Solving this equation for 0 and 180 msec. yields a value of 220 msec. which is higher than the control mean of Group E (213 msec.). However since some *Ss* still gave significantly longer RTs when S_2 appeared 180 msec. after S_1 and in view of the lower mean RT of Group C as compared with the control RTs of Group E, it is likely that there is still some effect of S_2 even as long as 180 msec. after the onset of S_1. Taking the curve as a whole we find that the maximal effects are obtained from 40 to 140 msec. following S_1 for it is in this region that the parabola is fairly flat. Variations in luminance, hue, or other characteristics of S_2 and/or S_1 would undoubtedly change the region of maximal effect.

The inhibiting effect of S_2 appears even more clearly when RTs of Groups E and C are plotted as a function of trials in Fig. 2. The wide separation between the curves for Group E when S_2 is given and Group C for all blocks of trials attests the strength of the inhibiting effect of S_2. Indeed, the two curves are more widely separated on the whole during the last 180 trials than during the first 180 trials. Analysis of variance showed these two blocks to be significantly different ($P < .05$). On the other hand, the plot of the trials in

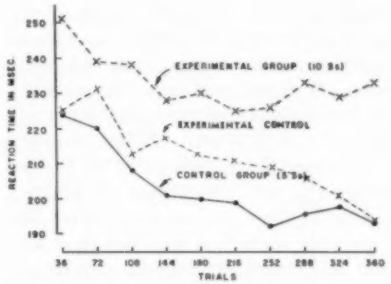


FIG. 2. Decrease in RT in Groups E and C as a result of practice. (Group C improves to a greater extent than does Group E when S_2 is present showing that repetition does not counteract the inhibiting effect of the interpolated stimulus, but the RTs of Group E when S_2 is omitted as shown by the "Experimental Control" curve, are identical with those of Group C after 360 trials.)

which only S_1 was given Group E (intermediate curve in Fig. 2) shows that with practice *Ss* in Group E were able to react as quickly to a single stimulus as were the *Ss* in Group C toward the end of the 360 trials. There is, therefore, a differential effect of practice in the single-stimulus condition and in the two-stimulus condition: the carry-over from double stimulation to single-stimulus condition is counteracted by practice, but practice does not overcome the inhibiting effect of S_2 in the two-stimulus condition. These findings point to nonattitudinal processes as the basis for the inhibiting effects of S_2. Finally, it should be noted that the inhibiting effect of S_2 is negated to some extent because of the drop in the Group E curve with the two-stimulus condition from Trial 36 to Trial 144 after which there seems to be no further practice effect.

Individual differences in the extent to which *Ss* are affected by the interpolated stimulus are striking. Group E divides into two subgroups, one of which is markedly and definitely affected by S_2 while the other, on the whole, is not. Overall *t* tests of differences between the experimental and control trials of the first subgroup were all significant at or beyond the .01 level; those for the other subgroup were not significant. When the 180 differences (18 intervals × 10 *Ss*) between the two conditions of stimulation are tested by means of correlated *t* tests for each *S*, it is found that 3 *Ss* had only two significant differences and 1 *S* had only three significant differences whereas the other 6 *Ss* had 11, 13, 14, 14, 16, and 17 significant differences out of a possible 18. It thus appears that some *Ss* are very much more affected by the interpolated stimulus than are others. Comments made by some of the *Ss* support this view. One *S*, for example, remarked again and again that "something was wrong with the apparatus" when the second stimulus appeared. These results do not imply that some *Ss* may be impervious to the effects of all secondary stimuli: Had the second stimulus been more intense than the first or had it been presented in another sense modality or otherwise made more impressive, it is highly probable (indeed certain, in our minds) that they would have given significantly longer RTs. Even these 4 *Ss*, as pointed out above, did have significantly longer RTs on some of the interpolated intervals.

Discussion

The data of this experiment seem to establish very clearly that a stimulus following the primary signal to react has the effect of lengthening RT according to a parabolic function, the effect being minimal 10 msec. after the onset of the primary stimulus, increasing to about 40 msec., and having maximal effect from 40 msec. to 140 msec. after which RT approaches the single-stimulus condition and is not statistically different from RT with one stimulus at the 180-msec. interval. The inhibiting effect of the second stimulus is found even after 360 repetitions although learning has clearly occurred in the single-stimulus condition as shown by the decline in the RT × Trials curve (Fig. 2) for both Group C and Group E. It appears unlikely that higher order, attitudinal, or other perceptual-cognitive processes can be responsible for the effect described here. Rather it seems more probable that lower order, automatically acting mechanisms not under voluntary control, are at work. The difficulty in explaining this phenomenon springs from the fact that a stimulus coming when the response is almost completed can exercise an inhibiting effect. At first sight it might appear that we are dealing with a higher order mechanism in that 4 of the 10 *Ss*

were not significantly affected by S_2 but we saw that even these *S*s had significantly longer RTs on 2 or 3 of the 18 intervals than on their control trials. We confidently expect that such *S*s would be influenced by a second stimulus if it were made more intense than the first or if a more compelling condition were used than was the case in this experiment.

An explanation, in physiological terms, at this time must be purely ad hoc and speculative. Two hypotheses, among the many we have considered, may be ventured as follows: (*a*) The first hypothesis involves the assumption of *inhibitory* fibers that are *faster* than the ordinary motor fibers innervating voluntary responses. Such fibers might be analogous to slow and fast afferent fibers that are known to function in mediating *slow* and *fast* pain or to differences in speed of afferent conduction associated with differences in fiber size. Whether or not such fast, inhibitory *efferent* fibers exist we do not know and if none have been discovered these data warrant a search for them by electrophysiological means. (*b*) The second hypothesis, while less suggestive of specific neurological or physiological mechanisms, has behavioral implications and may be stated as follows: Let us assume that the reflex arc is a total, ongoing process such that a disturbance in any part of it disrupts the ongoing activity with the result that a new integration is required for the activity to be resumed or completed. On this basis *S* is set to react to a single stimulus and when the second stimulus appears, even though he has not been instructed to react or attend to it, nevertheless it breaks into the ongoing response with consequent lengthened RT. It is evident that an hypothesis that explains the inhibiting effect of a second stimulus after a very short interval, e.g., 10–40 msec., may not explain its effect after a comparatively long interval, e.g., 170 msec., or vice versa. An hypothesis must explain lengthened RTs found with all intervals used in this study.

Summary

An earlier finding that a light following a sound or a sound following a light after an interval of 75 msec. lengthened RT to the stimulus first presented was verified in the present experiment in which the stimuli were both visual. RT to S_1 was significantly increased when S_2 followed S_1 at intervals ranging from 10 to 170 msec. with the maximum effect occurring from 40 to 140 msec. Trend analysis showed the quadratic components in the Intervals × RT data to be significant. While RT with S_2 present decreases with practice, the greater improvement of control *S*s as compared with experimental *S*s shows that the inhibitory effect of S_2 on S_1 is not completely negated by 360 repetitions. Four of the 10 experimental *S*s failed to give significantly lengthened RTs over all 18 intervals although they did have significantly higher RTs at some of the intervals. While the individual differences and results of practice seem to argue in favor of attitudinal factors as responsible for the effect, other facts argue against this explanation. Two hypotheses were discussed but neither seems completely satisfactory to explain all the facts.

References

Grant, D. A. Analysis-of-variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–154.

Teichner, W. H. Recent studies of simple reaction time. *Psychol. Bull.*, 1954, **51**, 128–149.

Todd, J. W. Reaction to multiple stimuli. *Arch. Psychol., N. Y.*, 1912, Whole No. 25.

(Received July 27, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 206–214

RESPONSE SUPPRESSION IN PERCEPTUAL DEFENSE[1]

ROBERT B. ZAJONC

University of Michigan

Recent theorizing maintains that the phenomenon of perceptual defense can be accounted for in terms of response processes. Elevated thresholds to taboo words are now generally regarded as reflecting a response bias deriving from either previous experience (Goldiamond & Hawkins, 1958), set (Postman, Bronson, & Gropper, 1953), or conflict (Brown, 1961), rather than a defensive perceptual blocking. Although the core of the issue deals with the relative contributions of the stimulus and response to the perceptual defense effect, studies attempting to evaluate such relative contributions have been rather few (Matthews & Wertheimer, 1958; Neisser, 1954). It is the purpose of this experiment to determine the extent to which both recognition threshold and the galvonic skin response (GSR) are influenced by the stimulus and to what extent by the response, using a procedure first suggested by Garner, Hake, and Eriksen (1956). First, threshold and GSR data for a set of taboo and neutral words were obtained by means of standard methods. Secondly, a paired-associate list was constructed using the previously exposed words as stimulus terms and a new set of taboo and neutral words as response terms. Some taboo stimuli were paired with taboo, others with neutral response terms. Neutral stimuli too, were sometimes paired with neutral and sometimes with taboo response terms. All *Ss* learned the paired-associate list to a criterion. The third step consisted of a repeated threshold and GSR assessment of the original stimuli. Now, however, one group was required to indicate recognition as before, i.e., by reading out loud the word presented tachistoscopically, and another by saying the appropriate response term which they have learned in the previous paired-associate task. Thus, the second group was given an opportunity to indicate recognition by means of responses whose emotional significance was either positively or negatively correlated with the emotional significance of the stimulus.

[1] This work was supported by Grants AS 49(638)-367 and G-4951 from the Air Force Office of Scientific Research and the National Science Foundation. I wish to thank James Taylor for his assistance in this study, and Dorwin Cartwright and Arthur Platz for reading the manuscript.

METHOD

Subjects.—Forty male *Ss*, all enrolled at the University of Michigan, participated in the experiment. They were randomly assigned to two experimental groups consisting of 20 *Ss* each. The *Ss* were paid $1.25 per hour for participation in the experiment.

Apparatus.—Gerbrands' transparent mirror tachistoscope with an instant start fluorescent lamp circuit was employed to present stimuli. Skin resistance changes were observed by means of a Lafayette psychogalvanometer Model 603-A.

Materials.—Stimulus words were printed in black 2-in. block letters and presented in the center of the exposure field on gray 10×12 in. cards (54.5% reflectance). Stimulus-response pairs in the paired-associated task were shown in the same manner. Twelve taboo and 12 neutral words were selected from McGinnies' (1949) original list of 18 words to which equivalent neutral and taboo stimuli were added. Half of the taboo and half of the neutral words were used as stimuli in the threshold assessment and as stimulus terms in the paired-associate training task. The remainder of the list was used as response

terms in the paired-associate task. Three taboo stimuli were paired with three taboo responses, three taboo stimuli were paired with three neutral responses, three neutral stimuli were paired with three taboo responses, and three neutral stimuli with three neutral responses. These sets of words will be referred to as the TT, TN, NT, and NN sets. The 24 words were APPLE, BROOM, CANDY, CHAIR, CHILD, FLOOR, MUSIC, RAINS, RIVER, SHELF, STOVE, TRADE, BALLS, BELLY, BLEED, FAIRY, FILTH, HYMEN, KOTEX, PENIS, PUBIC, RAPED, VOMIT, WHORE.

Procedure.—As briefly outlined above, the procedure consisted of two recognition threshold and GSR assessment sessions separated by an intervening paired-associate learning task. The *Ss* were divided into two groups of 20 *Ss* each, one of which was required during the second threshold assessment session to indicate recognition in terms of the stimuli presented (Group S), the other in terms of the response terms paired with the stimuli (Group R).

Thresholds were obtained by the ascending method of limits in .01-sec. steps beginning with .05 sec. below *S*'s threshold to a neutral training word. The intertrial intervals were approximately 30 sec. Some of the words were shown more than once in order to eliminate prerecognition guesses during the second threshold assessment, when full knowledge of the list was already available to *Ss*. For those words data from the first presentation alone were included in the analysis. The criterion of threshold was the first correct recognition of the word in Group S and the first emission of the correct response term in Group R.

The *S* was seated with his head against the eyepieces and with his hand to which electrodes were affixed lying relaxed on the table. A rest period of 1 min. was given after the first threshold assessment session and after the paired-associate training.

The tachistoscope was operated by an adult male and the psychogalvanometer by an adult female. The GSR readings were taken in terms of reduction in resistance from the basal resistance level, which was adjusted for each stimulus exposure. Only those reactions which occurred within 5 sec. following stimulus exposure, and only those for which the resistance returned to the immediate neighborhood of the pre-exposure level were recorded. GSR readings were taken on every presentation of the stimulus word. Since some *Ss* recognized the word on the fourth exposure, only two prerecognition trials and the recognition trial were considered. Thus, for each *S* three GSR scores were computed for each set of stimuli, and for the purposes of analysis all were converted to standard scores with a mean of 50 and *SD* of 10 for all 40 *Ss*.

TABLE 1

MEAN RECOGNITION THRESHOLDS (SEC.) BEFORE PAIRED-ASSOCIATE TRAINING

Group	Later PA Conditions				Words	
	TT	TN	NT	NN	All Taboo	All Neutral
S	.219	.228	.193	.198	.223	.196
R	.234	.233	.206	.213	.234	.210
Both	.227	.231	.200	.206	.229	.203

The paired-associate task was conducted using a 2-sec. interval for the presentation of the stimulus and a 2-sec. interval for the presentation of the pair, with 20 sec. between trials. All terms were presented tachistoscopically. The order of the stimuli was randomly altered from trial to trial. Three consecutive correct anticipations of the entire list were used as the criterion.

RESULTS

Recognition Threshold and GSR before Paired-Associate Training

Mean recognition thresholds obtained before paired-associate training are shown in Table 1. The analysis of variance for these results showed that the only significant effect is due to the difference between taboo and neutral words ($F = 17.08$, $P < .001$). Although the mean recognition thresholds for Group R are somewhat higher than those for Group S, this difference is not significant. Also, no significant differences were obtained between taboo words to be later used with taboo responses (TT) and taboo words to be later used with neutral responses (TN). Nor was there any difference between neutral words to be later used with taboo responses (NT) and neutral words to be later used with neutral responses (NN).

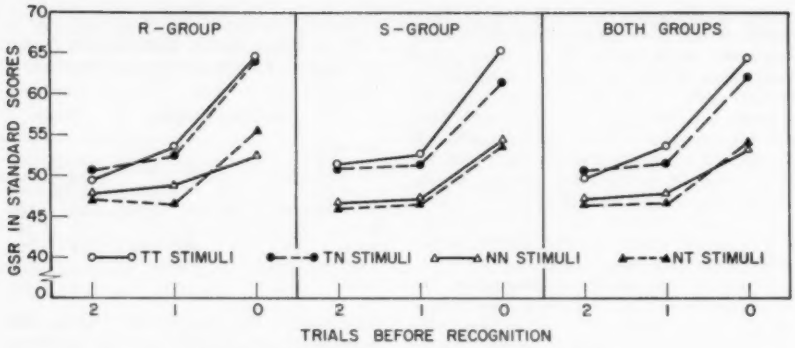


FIG. 1. GSR before paired-associate training.

The GSRs are shown in Fig. 1. Again no difference between the experimental groups was found. It is evident from the results that on all trials taboo words exceed neutral words in GSR ($F = 30.99$, $P < .001$). It is also clear that there is a considerable rise in the GSR on the recognition trial ($F = 32.94$, $P < .001$). No significant differences between TT and TN words as well as between NN and NT words were found for either of the two groups.

Paired-Associate Learning

Average number of trials to learn the four sets of associations are presented in Table 2. The means represent the number of trials which *S*s required to learn a given association to a criterion of three correct anticipations, averaged for the three items in each set. Shown in Table 2 is also the average number of errors for each set of pairs. The results indicate that the four sets of associations were not learned at the same rate (see Table 3). In particular, the TN pairs seem to be the most difficult, and the TT easiest. The analysis of variance presented in Table 3 shows a significant effect due to differences between word sets, which is primarily due to the type of response. In general, pairs with a taboo response require fewer trials and lead to fewer errors than pairs with neutral responses. Of particular importance to the present experiment is the difference between the TN and NT pairs. If speed of learning and number of errors reflect the degree to which a given response has become attached to the stimulus word, the TN stimuli should, during the subsequent threshold and GSR assessment of Group R, be more handicapped than NT stimuli. The differences between these pairs in both trials to criterion and average number of errors are significant at the .001 level.

Recognition Threshold and GSR after Paired-Associate Training

Group R.—Table 4 shows recognition thresholds for the four sets of words for the condition in which *S*s indicated recognition by means of the response term acquired during the paired-associate training. It is apparent that, compared with those obtained before the paired-associate training, the thresholds to all the words are considerably reduced. It is

TABLE 2

MEAN TRIALS AND ERRORS TO CRITERION IN PAIRED-ASSOCIATE LEARNING

Group	Pairs				Words			
	TT	TN	NT	NN	Taboo Stimuli	Neutral Stimuli	Taboo Responses	Neutral Responses
Group S								
Trials	4.56	5.62	4.37	5.13	5.09	4.75	4.47	5.38
Errors	4.55	7.10	3.70	6.10	5.83	4.90	4.13	6.60
Group R								
Trials	4.21	4.95	4.48	4.28	4.58	4.38	4.35	4.62
Errors	2.90	5.00	3.80	3.55	3.95	3.68	3.35	4.28
Both								
Trials	4.39	5.29	4.43	4.71	4.84	4.57	4.41	5.00
Errors	3.73	6.05	3.75	4.83	4.89	4.29	3.74	5.44

also clear that no longer does the recognition threshold totally depend on the stimulus. There is a considerable effect due to the response which *S* utilizes in indicating recognition. It should be pointed out that *S*'s ability to give evidence of recognition, not by means of a word which is presented but by means of a response previously learned, depends on the degree to which these responses were fixated. It will be recalled that the learning of the four types of associations was not uniform. In particular there was a considerable difference between the TN and the NT pairs, in favor of the latter. Moreover, the examination of the results on paired-associate learning disclosed a significant effect due to individual differences. The *F* ratios evaluating the individual difference effect were 8.79 for trials to criterion, and 11.41 for errors, which for the degrees of free-

TABLE 3

ANALYSIS OF VARIANCE FOR DATA IN TABLE 2

Source	*df*	Trials to Criterion		Errors	
		MS	*F*	*MS*	*F*
Treatments (A)	1	8.06	.83	96.09	1.09
Words	3	6.69	6.03***	48.54	6.23***
S component (B_1)	1	3.11	3.27	14.40	2.68
R component (B_2)	1	13.40	8.93**	115.60	11.73**
$B_1 \times B_2$	1	3.57	4.12	15.62	1.92
Treatments × Words	3	1.78	1.60	13.45	1.73
$A \times B_1$	1	.24	.25	4.23	.79
$A \times B_2$	1	3.93	2.62	24.03	2.44
$A \times B_1 \times B_2$	1	1.18	1.37	12.10	1.49
Error (b)	38	9.76		88.91	
Error (w)	114	1.11		7.79	
$Ss \times B_1$	38	.95		5.38	
$Ss \times B_2$	38	1.50		9.85	
$Ss \times B_1 \times B_2$	38	.86		8.13	

** $P = .01$.
*** $P = .001$.

TABLE 4

MEAN RECOGNITION THRESHOLDS (SEC.) IN GROUP R AFTER PAIRED-ASSOCIATE LEARNING

Group	PA Pairs				Words			
	TT	TN	NT	NN	Taboo Stimuli	Neutral Stimuli	Taboo Responses	Neutral Responses
Rapid learners	.151	.138	.145	.138	.145	.142	.148	.138
Slow learners	.178	.180	.174	.172	.179	.173	.176	.176
All *Ss*	.165	.159	.160	.155	.162	.158	.163	.157
Adjusted means for all *Ss*	.168	.153	.160	.157	.161	.159	.164	.155

dom given are significant well beyond the .001 level. We would expect a more reliable test of the relative contributions of the stimulus shown and of the response given from *Ss* who learned these responses well. Group R was therefore divided at the median number of trials to criterion, and the recognition thresholds for the rapid and slow learners are shown in Table 4, and the analysis of variance in Table 5. It is clear from Table 4 that slow learners manifest considerably higher recognition thresholds for *all* the words. The difference between groups is significant at better than the .05 level. It appears that the slow learner's recognition threshold depends primarily on the type of stimulus presented, while that of rapid learners on the response which they were required to make. However, the Groups × Stimulus × Response interaction was not significant. The overall results, however, indicate that the effects due to the stimulus component were not significant while those due to the response were significant. Further support for the conclusion that recognition threshold depends primarily on the type of response required is obtained when the data are adjusted for differences in learning the four types of associations. The mean recognition thresholds, adjusted by means of the regression equation relating the former to the number of trials, are shown at the bottom of Table 4. Analysis of covariance performed on these results disclosed a significant effect due to the response component ($F = 8.78$, $df = 1/17$) and no effects due to stimulus.

TABLE 5

ANALYSIS OF VARIANCE FOR DATA IN TABLE 4

Source	*df*	*MS*	*F*
Groups (Rapid vs. Slow) (A)	1	24,945	6.72*
Words	3	338	2.54
S component (B_1)	1	466	2.13
R component (B_2)	1	546	5.00*
$B_1 \times B_2$	1	1	<1.00
Groups × Words	3	131	<1.00
$A \times B_1$	1	35	<1.00
$A \times B_2$	1	536	4.91*
$A \times B_1 \times B_2$	1	122	1.67
Error (b)	18	3,711	
Error (w)	54	133	
$Ss \times B_1$	18	218	
$Ss \times B_2$	18	109	
$Ss \times B_1 \times B_2$	18	73	

* $P = .05$.

The GSR data shown in Fig. 2 follow a similar pattern. Again, as compared with the results obtained before paired-associate training, the GSRs are weaker. The analysis of variance (Table 6) shows a significant effect due to the differences between

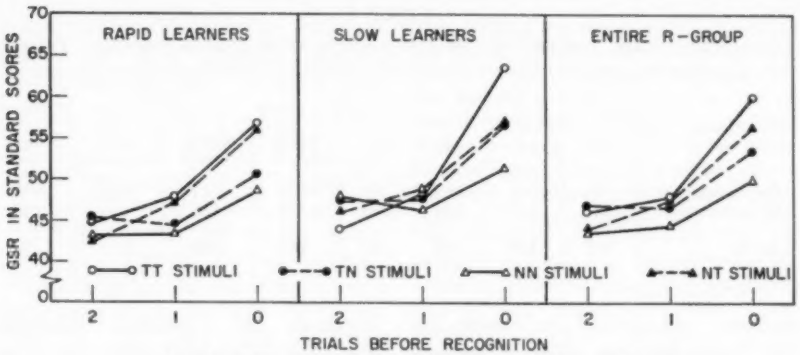


FIG. 2. GSR after paired-associate training (Group R).

words which seems to be a function of the stimulus and of the response component as well. The results of both groups combined indicate that on the second prerecognition trial the GSRs do not follow any particular

TABLE 6

ANALYSES OF VARIANCE FOR DATA IN FIG. 2 AND 3

Source	*df*	Group R		Group S	
		MS	*F*	*MS*	*F*
Groups (Rapid vs. Slow) (A)	1	522.4	1.19	161.7	<1.00
Words	3	171.1	6.35**	78.6	3.48*
S component (B_1)	1	149.6	7.03*	227.0	7.73*
R component (B_2)	1	356.5	10.57**	2.5	<1.00
$B_1 \times B_2$	1	7.3	<1.00	6.3	<1.00
Trials (C)	2	2,235.6	28.97***	942.5	12.57***
Groups × Words	3	9.6	<1.00	26.1	1.16
$A \times B_1$	1	.3	<1.00	16.9	<1.00
$A \times B_2$	1	27.8	<1.00	51.4	1.92
$A \times B_1 \times B_2$	1	.8	<1.00	10.1	<1.00
Words × Trials	6	136.7	5.44***	38.6	3.08**
$B_1 \times C$	2	87.8	4.27*	79.0	4.71*
$B_2 \times C$	2	319.6	12.47***	6.6	<1.00
$B_1 \times B_2 \times C$	2	2.8	<1.00	30.4	4.94*
Groups × Trials	2	25.3	<1.00	39.4	<1.00
Groups × Words × Trials	6	32.6	1.30	22.2	1.77
$A \times B_1 \times C$	2	77.5	3.77*	36.7	2.19
$A \times B_2 \times C$	2	5.7	<1.00	4.8	<1.00
$A \times B_1 \times B_2 \times C$	2	14.7	<1.00	25.1	4.09*
Error (b)	18	439.1		441.1	
Error $(w)_1$: *Ss* × Words	54	27.0		22.6	
Ss × B_1	18	21.3		29.4	
Ss × B_2	18	33.7		26.8	
Ss × $B_1 \times B_2$	18	25.9		11.6	
Error $(w)_2$: *Ss* × Word × Trials	36	77.2		75.0	
Error $(w)_3$: *Ss* × Words × Trials	108	25.1		12.6	
Ss × $B_1 \times C$	36	20.6		16.8	
Ss × $B_2 \times C$	36	25.6		14.7	
Ss × $B_1 \times B_2 \times C$	36	29.2		6.2	

* $P < .05$.
** $P < .01$.
*** $P < .001$.

TABLE 7

MEAN RECOGNITION THRESHOLDS (SEC.) IN GROUPS AFTER PAIRED-ASSOCIATE TRAINING

Group	PA Pairs				Words			
	TT	TN	NT	NN	Taboo Stimuli	Neutral Stimuli	Taboo Responses	Neutral Responses
Rapid learners	.153	.151	.143	.141	.152	.142	.148	.146
Slow learners	.167	.175	.154	.162	.171	.158	.160	.168
All *Ss*	.160	.163	.149	.152	.161	.150	.154	.157

pattern. However, the curves for the rapid learners show a pattern of particular interest. On the second prerecognition trial the GSRs seem to depend primarily on the stimulus component; their order is TN, TT, NN, and NT. As the *Ss* approach recognition the stimulus effect is gradually replaced by the response effect and the GSRs are ordered according to the response. One may interpret this result to mean that stimulation present two trials before recognition is probably too weak to call out strong anticipatory partial responses. As soon as the stimulation gains in strength and becomes capable of evoking some parts of the learned response, the autonomic reactions lose their dependence upon the stimulus and begin to be dominated by the response component.

TABLE 8

ANALYSIS OF VARIANCE FOR DATA IN TABLE 7

Source	*df*	*MS*	*F*
Groups (Rapid vs. Slow Learners) (A)	1	10,160	3.12
Words	3	896	3.03
S component (B_1)	1	2,532	5.34*
R component (B_2)	1	157	<1.00
$B_1 \times B_2$	1	0	<1.00
Groups × Words	3	150	<1.00
$A \times B_1$	1	131	<1.00
$A \times B_2$	1	419	1.68
$A \times B_1 \times B_2$	1	1	<1.00
Error (b)	18	3,088	
Error (w)	54	296	
Ss × B_1	18	474	
Ss × B_2	18	250	
Ss × $B_1 \times B_2$	18	164	

* $P = .05$.

The mean GSR reaction for rapid learners was 47.55 and for slow learners 50.53, but as is evident from Table 6 this difference was not significant.

Group S.—The principal purpose of the paired-associate learning task was to enable *Ss* to give evidence of recognition of the stimulus words without having to say them. However, it is possible to argue that the training simultaneously produced temporary changes in the emotional quality of the stimulus words. Thus, taboo stimuli which were paired with neutral responses could, by virtue of the repeatedly reinforced association, have become emotionally "neutralized." Similarly, conditioning a taboo response to a neutral stimulus word might have affected the emotional quality of the latter. These eventualities are of course quite remote because of the small number of conditioning trials involved. If conditioning of the type suggested has in fact taken place then the recognition thresholds and the GSR data should show the same patterns in Groups S and R. The average recognition thresholds for Group S are shown in Table 7, and the analysis of variance for these results has shown

effects only due to the stimulus component. It is of interest to note that as was the case in Group R slow learners in Group S also showed somewhat higher recognition thresholds than rapid learners. However, this difference failed to reach an acceptable level of significance.

Neither do the GSR results shown in Fig. 3 suggest any conditioning effect. Besides the increase in reactions over trials, the only significant effect is that due to the stimulus component. The analysis of variance in Table 8 shows an F ratio significant at the .05 level for the stimulus component. On the trials immediately preceding recognition there is a slight but not significant response effect for rapid learners. Also, as observed before, the GSRs of rapid learners are somewhat less than those of slow learners (45.69 and 47.16, respectively), but this difference is decidedly not reliable.

Discussion

The evidence presented failed to disclose perceptual effects of any significance. The recognition threshold was found to be a function not of what S saw but what he had to say. Moreover, GSR data follow an identical pattern. The GSRs were found to be produced not by the stimulus alone, but depended primarily on the response required of S. The results are best accounted for by Brown's (1961) competing response theory. Irrespective of the stimulus, if the responses were in conflict with an inhibitory tendency, that is, if S had to make a vulgar response, both recognition threshold and GSR were elevated. Stimuli arousing no response conflict failed to produce differential thresholds and GSRs irrespective of their "emotionality." Further support for the response competition hypothesis is seen in the GSR data. In general, the differences in the GSRs were found to increase over trials, reaching their peak upon recognition. To the extent that the GSRs reflect response conflict, one would expect that with increasing exposure time both the positive and the negative tendencies increase, thus generating a stronger conflict.

There is evidence in the data that recognition threshold and GSR are also subject to variation as a result of not only a conflict between a positive and negative tendency, but also as a result of a conflict between competing excitatory tendencies. First we note that both are markedly reduced after familiarization with the stimuli. Before paired-associate learning the response alternatives available to Ss are many, and all of these are in competition. The training

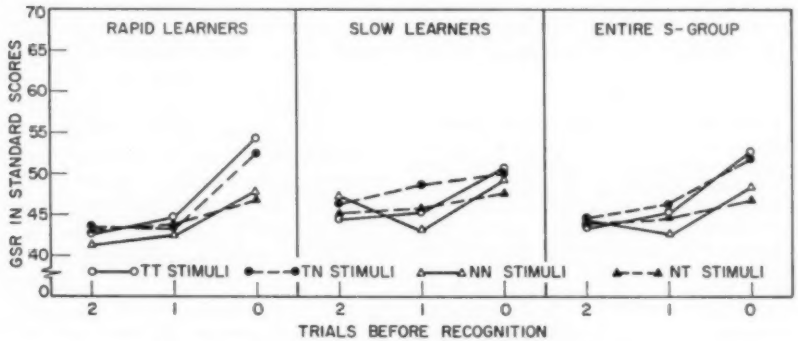


Fig. 3. GSR after paired-associate training (Group S).

reduces them to 12, thus reducing the extent of response competition involved. Secondly, consistent differences in the overall recognition threshold and GSR reactions between the rapid and slow learner were found. If one views the speed of the paired-associate learning and the mean number of errors as an index of the amount of response competition present, these results become quite meaningful.

It is not claimed here that the perceptual defense phenomenon has been disproven. But if the phenomenon is empirically demonstrable its proof must be established by experimental methods other than those commonly used. Perhaps Blum's (1954) forced-choice technique of threshold assessment holds best promise since it eliminates possible effects due to the response process.

Summary

The role of stimuli and responses in perceptual defense was examined by first obtaining recognition thresholds and GSRs to taboo and neutral words. Subsequently, *Ss* learned a paired-associate list with the original words serving as stimulus terms and a new set of words as response terms. Half of the neutral stimuli were paired with neutral and half with taboo responses. The same was true of taboo stimuli. Following training, recognition thresholds and GSRs were again measured with one group required to indicate recognition by means of response terms and another by means of stimulus terms. Both recognition threshold and GSR were found to depend primarily on the response required of the *Ss* in indicating recognition.

References

Blum, G. S. An experimental reunion of psychoanalytic theory with perceptual vigilance and defense. *J. abnorm. soc. Psychol.*, 1954, **13**, 94–99.

Brown, J. S. *The motivation of behavior.* New York: McGraw-Hill, 1961.

Garner, W. R., Hake, H. W., & Eriksen, C. W. Operationism and the concept of perception. *Psychol. Rev.*, 1956, **63**, 149–159.

Goldiamond, I., & Hawkins, W. F. Vexierversuch: The log relationship between word frequency and recognition obtained in the absence of stimulus words. *J. exp. Psychol.*, 1958, **56**, 457–463.

Mathews, A., & Wertheimer, M. A "pure" measure of perceptual defense uncontaminated by response suppression. *J. abnorm. soc. Psychol.*, 1958, **57**, 373–376.

McGinnies, E. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, **56**, 244–251.

Neisser, U. An experimental distinction between perceptual process and verbal response. *J. exp. Psychol.*, 1954, **47**, 399–402.

Postman, L., Bronson, W. C., & Gropper, G. L. Is there a mechanism of perceptual defense? *J. abnorm. soc. Psychol.*, 1953, **48**, 215–224.

(Early publication received January 5, 1962)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 215–226

FACTORS IN THE RETENTION AND RELEARNING OF PERCEPTUAL-MOTOR SKILL [1]

EDWIN A. FLEISHMAN

Yale University

AND JAMES F. PARKER, JR.

Psychological Research Associates, Arlington, Virginia

In several previous reports, Parker and Fleishman have described studies of complex tracking performance. The first of these (Parker & Fleishman, 1959, 1960) attempted to predict performance at different stages of learning a complex tracking skill. Special efforts were made to predict high levels of proficiency after extensive practice (17 sessions distributed over 6 weeks) with this task. A second study (Parker & Fleishman, 1961) made use of information about the components of tracking skill to facilitate the learning of this skill. The present study is an investigation of factors in the retention and relearning of this same skill.

Previous studies of motor skill retention (e.g., Ammons, Farr, Bloch, Neumann, Dey, Marion, & Ammons, 1958; Battig, Nagel, Voss, & Brogden, 1957; Bell, 1950; Jahnke, 1958; Jones & Bilodeau, 1953; Leavitt & Schlosberg, 1944; Mengelkoch, Adams, & Gainer, 1958; Reynolds & Bilodeau, 1952) all present evidence that continuous control, perceptual-motor skills are well retained over fairly long periods of no practice. What loss occurs appears to be quickly regained. The present study is a more comprehensive study of factors in retention and relearning, using a highly complex continuous control task requiring considerable practice for initial learning.

[1] This study was performed under Contract Nonr 3065(00) between Psychological Research Associates, Incorporated, and the Office of Naval Research.

While essentially a laboratory study, the task was designed to simulate a complex skill, i.e., that of a pilot flying a radar intercept mission. The problem of retention over extended periods without practice is especially critical here. Furthermore, where the skills are of such complexity, the problem of finding the optimum conditions for retraining becomes even more critical.

Specifically, the following questions were investigated. How well is such a skill retained without practice? What is the relation between the length of the "no practice" interval and level of retention? If there is a loss in proficiency, how much practice is required to regain proficiency? What is the relation between retention and level of proficiency after the original learning? Is the *type* of initial training related to retention? What is the relative effectiveness of a distributed vs. massed retraining schedule? Does the type of retraining schedule affect later performance as well as performance during retraining?

METHOD

Task.—The criterion task consisted of a tracking device constructed so as to simulate roughly the display characteristics and control requirements of an air-borne radar intercept mission. The task of *S* was to maintain the target dot at the center of the oscillograph display, while at the same time nulling a sideslip indicator. That is, *S* envisioned himself to be flying the attack phase of an air-borne rador intercept mission. Thus, if the target was to the right, *S* made appropriate control movements to steer the craft to the right. These movements would bring him

on target and the dot would return to the center of the display. All turning movements required coordinated action of stick and rudder controls.

Three identical tracking devices were constructed especially for purposes of this study. Photographs and complete schematics of all components are presented elsewhere (Parker & Fleishman, 1959). These devices and related scoring consoles allowed for the testing of from 1 to 3 *S*s simultaneously under the control of a single test administrator.

The *S*'s instrument panel contained two displays. The first consisted of a target dot presented on a cathode ray oscillograph. The target course was generated by setting the equation of a swinging pendulum into an analog computer. This produced a sine wave with a frequency of approximately 6 cycles per min. in the horizontal coordinate as the target course. The rate of decay in amplitude was approximately 5% per cycle. However, the dynamic characteristics of the overall task were such that *S*s neither perceived this drop in amplitude nor the fact that the target was programed in one coordinate only. Any control imbalances resulted in dot excursions of considerably larger magnitude than those provided by target programing.

Beneath the oscillograph was an inverted 3-in., zero-center voltmeter termed a "sideslip indicator." This meter reading indicated to *S* a "lack of coordination" in control actions when centering the target dot. This indicator did not constitute an independent task but rather provided *S* with additional information related to the primary task.

In performing this task *S*s used a standard aircraft control system involving a control stick and rudder pedals. These controls were coupled in a manner similar to those of an actual aircraft. Thus, application of right control stick pressure without proper amount of right rudder produced a sideslip to the left and a consequent left deflection on the sideslip indicator.

Control of the target dot in elevation was accomplished by forward and backward movements of the control stick. This was a pure second-order system resulting from the use of two cascaded integrators in the linkage between the control and display. Thus the acceleration of the target dot was directly proportional to stick displacement.

Control of the target dot in azimuth was accomplished by right and left movements of the control stick. This comprised a system involving acceleration control plus an exponential lag network. This dimension was mechanized, using three cascaded integrators with a negative feedback loop around one. The time constant of this lag network is 1 sec., i.e., it requires 1 sec. to achieve approximately two-thirds $(1-1/e)$ of the final signal resulting from a given stick displacement.

Control of the target dot in azimuth (envisioned as turning the aircraft) and centering the sideslip indicator (coordination display) were both affected by rudder pedal displacement as well as by the control stick. This rudder control of the sideslip indicator involves a simple lag network. Thus, the sideslip indicator displacement was directly proportional to rudder pedal displacement. Movement of the sideslip indicator by stick action represents a velocity control operating through two exponential lag networks.

Movement of the target dot by rudder pedal displacement approximates a pure velocity control.

Scoring.—The primary score was the *integrated absolute error score.* This was recorded at the conclusion of every trial and was produced by summing algebraically the three absolute error part scores in accordance with this relationship: $T=1/2X+1/2Y+Z$, where T = integrated absolute error score, X = absolute azimuth error, Y = absolute elevation error, and Z = sideslip (lack of coordination) error.

Initial training.—The initial learning data upon which this study is based were gathered during the course of two previous studies.[2] In each of these studies *S*s spent a total of 17 sessions, distributed over 6 weeks, mastering the tracking task, with each session consisting of 21 1-min. trials.

In the first study (Parker & Fleishman, 1959, 1960) no formal training instructions were administered. Although provided knowledge of results, *S*s learned the tracking task "on their own" with the single exception that any questions were answered. This group of *S*s will be referred to as Group I.

In the second study (Parker & Fleishman, 1961) 60 *S*s, referred to as Group II, spent an identical period of time mastering the same tracking task. These *S*s were administered a "common sense" training program. This consisted of an initial explanation and demonstration of the tracking device. This was followed by three sessions of practice with an *E* monitoring the entire operation and assisting *S* as required. The remaining 14 sessions consisted of individual practice

[2] The first study was concerned with correlates of performance at different stages of learning this task. The second study evaluated an experimental training program. Group II, in the present study, was the control group in that experimental training program.

followed by critiques with each *S* after Sessions 7, 11, and 15. As would be expected, terminal proficiency for the group trained under this program was significantly superior to that of the group which had no formal training.

Retention testing.—Seven groups of 10 *Ss* were brought back for retraining following various intervals of no practice. Intervals since training for Group II were 1, 5, 9, and 14 mo. Intervals since training for Group I were 9, 14, and 24 mo. The fact that it was possible to study the 9- and 14-mo. intervals for both groups allows a comparison of retention for the same intervals for two types of original training.

Each retention group was split into two subgroups of 5 *Ss* each. One subgroup was retrained during four intensive, continuous retraining sessions[3] during the same day. The other subgroup was retrained during four sessions, each scheduled 1 day apart. The purpose of this experimental breakdown was to allow an evaluation of the relative effectiveness of these two retraining schedules where one involved massed and the other distributed practice.

One week following the final retraining session all *Ss* were again tested for one additional session. The purpose of this additional testing was to allow a more adequate evaluation of the two types of retraining. This fifth retraining session was included to show whether any differences which might be found occurred only during the retraining program or whether these differences persisted in later performance. If transfer to later performance could be demonstrated, this could be attributed to differential *learning* during the course of retraining rather than to temporary performance factors (e.g., fatigue, inhibition) during the retraining.

Matching of retention groups.—It will be recalled that the seven retention groups were drawn from two groups of original trainees. One group (Group I) had been trained without benefit of specific guidance while the other group (Group II) had been trained with supplementary instruction and guidance. Accordingly, it was found (Parker & Fleishman, 1961) that Group II was superior in terminal proficiency although the number of practice sessions was identical for each group. For the present retention study an attempt was made to match the different retention samples on the basis of final performance level during original learning. This was done separately for the three retention samples drawn from Group I and for the four retention samples drawn from Group II.

[3] As before, a "session" includes 21 1-min. trials.

TABLE 1

FINAL PERFORMANCE LEVELS OF THE SEVEN RETENTION SAMPLES AFTER INITIAL LEARNING

Retention Interval (Months)	Group I No Formal Guidance			Group II Formal Guidance		
	N	Mean[a]	*SD*	*N*	Mean[a]	*SD*
1				10	246	82
5				10	232	57
9	9	294	119	9	232	59
14	7	263	88	9	230	71
24	8	260	117			

[a] Integrated absolute error score.

Scores (integrated error) attained by each *S* during the final session of original training were used as a basis for matching. These were converted to standard scores (stanines). An attempt was made to assign a proportionate number from each stanine level to each retention group in order to obtain a normal distribution representative of the original learning population. It soon became apparent that certain *Ss* needed to fulfill these requirements were not available for retesting. However, this procedure was followed as closely as possible. A preliminary analysis of variance for the first four samples tested indicated that they were not homogeneous. An adjustment was made by eliminating a few *Ss* whose final scores during the original training were extremely poor. This left a total of 62 *Ss* in the seven retention samples. Table 1 indicates the number of *Ss* in each retention group and the means and *SDs* of their final scores after initial learning. It can be seen that within each of the original learning groups adequate matching was achieved. Analyses of variance performed for each original training group confirmed that the retention samples could be considered comparable (Group I: $F = .22$, $df = 2/21$; Group II: $F = .09$, $df = 3/34$).

The comparability of these retention groups becomes especially apparent when one considers the range of *possible* scores from early to late learning. The curves in Fig. 1 and 2, for example, illustrate this.

RESULTS

Magnitude of retention.—A primary concern is the extent to which the developed performance capability deteriorates through time. However,

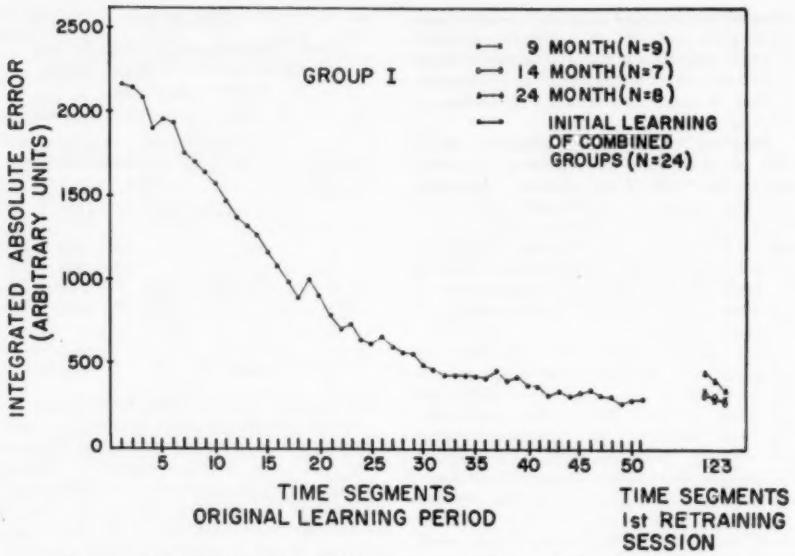


FIG. 1. Performance of Group I (no formal training) during original learning and following varying periods without practice.

a simple comparison of an *S*'s score following some period of no practice with his final score in initial training will not provide a complete understanding of performance loss. One must have information concerning the course of initial learning and the extensiveness of the training required to develop the skill. Figures 1 and 2 present, for Groups I and II, curves illustrating the course of initial learning followed by the results of the first session retention measurements. Each point in the initial learning curve is based on the average of the combined *S*s of the retention samples. Points on the abscissa are directly comparable for original training and retention and represent 6-min. periods within a practice session. Each session consisted of 21 min. of practice. The first three 1-min. trials were not scored to remove warm-up effects.

One of the primary conclusions drawn from Fig. 1 and 2 is that by comparison with the original learning of this skill there is little decrement in performance even for no practice periods of up to 24 mo. There is obviously somewhat more decrement in the 24-mo. group but recovery is rapid even during this first 21-min. retraining session.

Retention and length of interval.—In order to obtain a more precise description of performance at the beginning of retraining, the results of the first retraining session were plotted on a trial-by-trial (minute-by-minute) basis. Figure 3 shows that for Group I (no formal training) the major part of performance capability was regained during the first 2 or 3 min. of retraining. It can also be seen that the 24-mo. group is consistently poorer during this first retraining session than are the other two groups. However, even this group is improving to the level of the other groups within this first 21-min. retraining session. A much smaller loss occurs

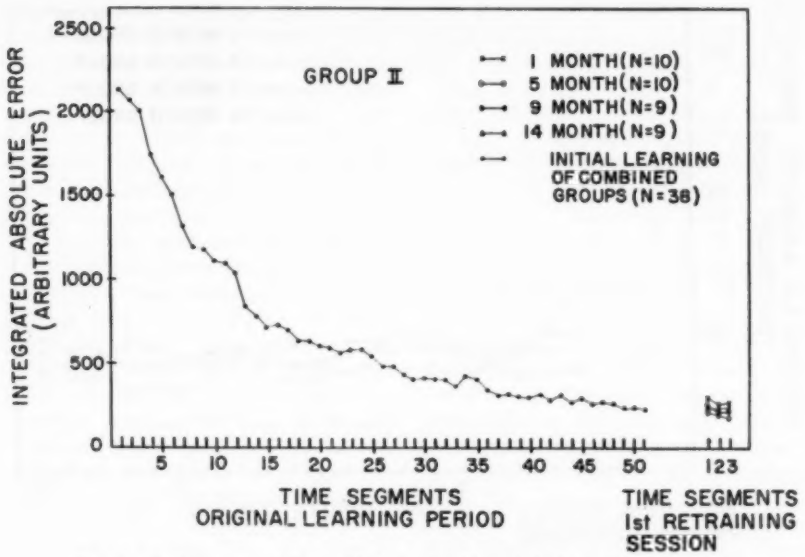


FIG. 2. Performance of Group II (formal guidance) during original learning and following varying periods without practice.

for the 9- and 14-mo. intervals and this is regained in just a few minutes. Differences between these latter two groups are negligible.

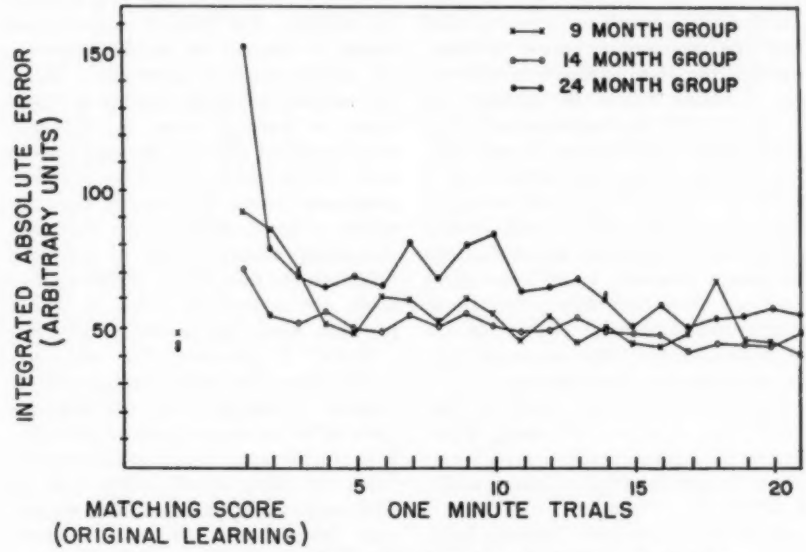


FIG. 3. Performance of Group I (no formal training) subgroups during the first retraining session.

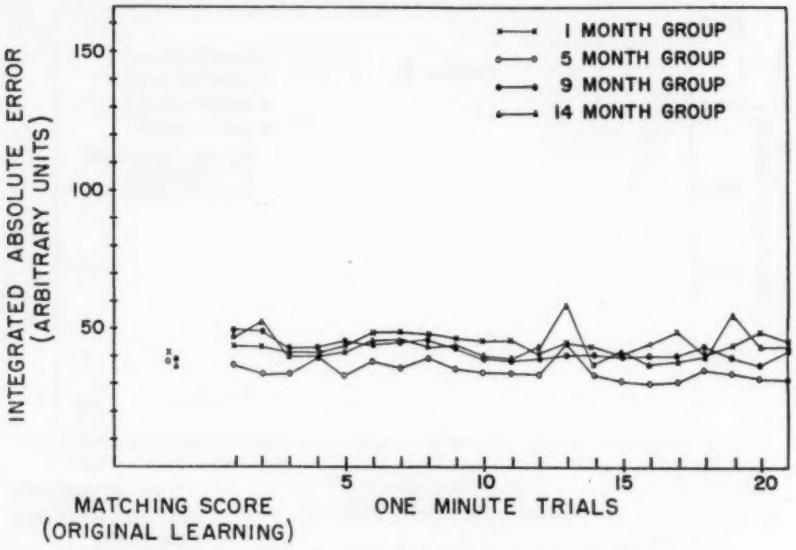


FIG. 4. Performance of Group II (formal guidance) subgroups during the first retraining session

Figure 4 indicates that Group II (formal guidance procedures) showed practically no deterioration without practice. It should be kept in mind that the maximum period without practice for this group was 14 mo. Our findings with this group are consistent with the findings for Group I in finding no differences in retention level for periods of no practice of 9 to 14 mo. The findings with Group II also indicate that the 9- and 14-mo. groups show no greater losses than do the groups with only 1 and 5 mo. of no practice. It is also shown that these groups, which exhibit essentially no forgetting, do not improve much during this first retraining session.

Figure 5, which is based on the first 1-min. trial of retraining, illustrates no performance loss as a function of longer intervals of no practice up to 14 mo.

Retention and original learning level. —Next an examination was made of the correlation between final performance level at the conclusion of the original learning period and performance level during the first retraining session. For this, an attempt was made to obtain as stable measures of performance as possible. Thus, the *original learning* measure is based upon an average score for the last three original practice sessions, or 44 min. of performance. The *retention score* represents the entire 18 min. which were scored during the first retraining session. (As in previous analyses, the first 3 min. of this session were not scored in order to avoid possible need for warm-up effects.)

Table 2 presents the obtained correlations between final level of original learning and performance after different intervals of no practice. It is readily apparent that all correlations are exceptionally high (.80 to .98), and all are statistically significant beyond the .01 level. Thus, there is virtually no change in the ordering of *S*s in any group with the

passage of time without practice. In order to obtain a single estimate of the relationship between retention and original learning level, all these cases were pooled together with scores of 40 other additional *Ss* not shown in Table 2. These last 40 *Ss* were *Ss* who learned initially under Group I procedures (no formal training) and who were brought back on a random basis for a single retraining session. Their retention intervals ranged from 9 to 25 mo. For this combined group ($N = 109$), having as it did a wide range of no practice intervals, the zero-order correlation between original learning level and retention score was found to be .80. A partial correlation coefficient between original learning and retention, with the effect of retention interval held constant, was .79. The loss of one point can be attributed to rounding error in the computational process.

The zero-order correlation between retention interval and retention score, for this combined sample of 109 *Ss* was .30; when initial learning is partialed out this drops to .23. This underscores the small amount of variance in the retention score attributable to the retention interval relative to the large amount of variance in retention due to initial learning level.

TABLE 2

CORRELATIONS BETWEEN ORIGINAL LEARNING LEVEL AND RETENTION TEST PERFORMANCE

Group	Retention Interval (Months)				
	1	5	9	14	24
I	—	—	.84	.93	.80
II	.89	.85	.98	.93	—

Note.—All *r*'s are based on *N*s of 10, except for Group I, 9-mo. entry, where *N* is 9.

An important question is whether the effect of initial learning upon retention performance is more important following short periods of no practice as opposed to longer intervals. In other words does the relation between initial learning level and retention level dissipate through time?

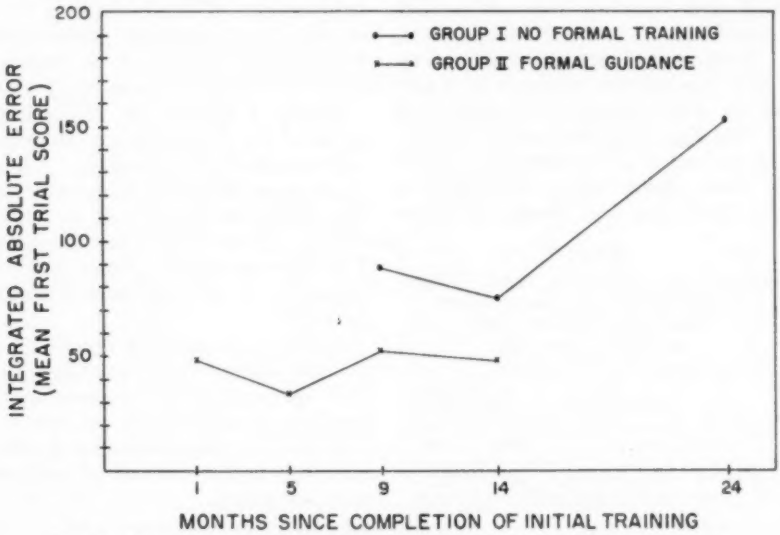


FIG. 5. The effect of retention interval upon first-trial retraining performance.

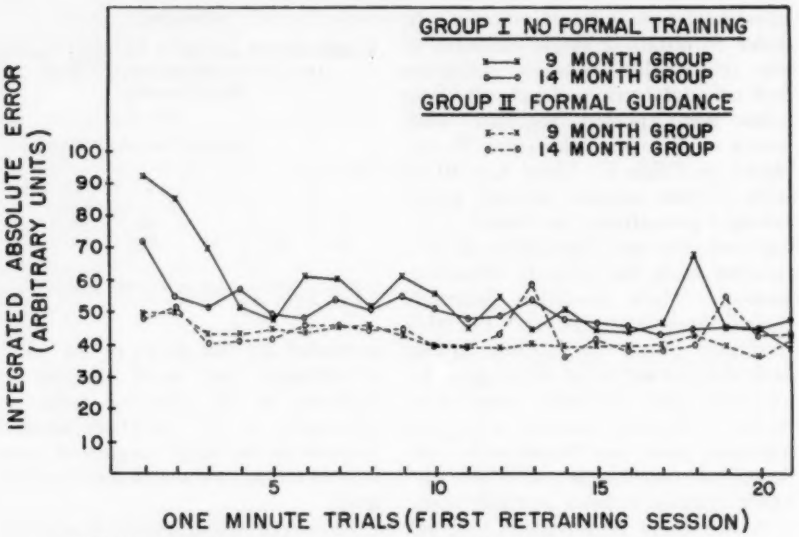


FIG. 6. Comparison of groups with same retention interval, illustrating the importance of prior proficiency level.

The correlations presented in Table 2 offer no evidence in support of this. The relationship between original learning level and retention performance appears to remain relatively high and constant through periods from 1 mo. to 24 mo. of no practice.

Retention and type of initial training.—The no-practice intervals of 9 and 14 mo. were common to Groups I and II. In Fig. 5 the differences between the two retention performance curves at these points reflect differences in the final learning levels of these groups. As described earlier, the differences in final learning level result from two types of initial training procedures. It still remains to be shown if the type of initial training, independent of final learning level, is related to retention performance. An analysis was made to separate the contribution of these two factors.

Figure 5 shows the performance levels of the 9- and 14-mo. *Ss* during the first minute of retraining. Figure 6 compares their performance during the entire 21-min. initial retraining session. It can be seen that the Group I *Ss* are consistently poorer than the Group II *Ss* with the same retention intervals.

Table 3 presents the results of an analysis designed to answer the question concerning the importance of type of initial training vs. level of proficiency at the conclusion of initial training, as determiners of performance retention. It was possible to select *Ss* from Groups I and II who were matched in terms of retention

TABLE 3

ANALYSIS OF RETENTION SCORES (INTEGRATED ABSOLUTE ERROR) TO DETERMINE IMPORTANCE OF TYPE OF INITIAL TRAINING

Group	*N*	Final Session: Initial Training		First Session: Retention	
		Mean	*SD*	Mean	*SD*
I	10	743.3	198.11	857.2	237.13
II	10	749.2	206.20	825.9	248.64

Note.—Subjects matched on retention interval and proficiency at conclusion of initial training.

interval (9 or 14 mo.) as well as in terminal proficiency at the conclusion of initial training. Table 3 presents the matching scores (terminal proficiency after initial training) and the mean scores obtained during the first retention session. No significant difference was found between scores ($t = .32, df = 9$). This indicates that the differences among our Groups I and II retention samples following periods of no practice of 9 and 14 mo. are a function of *level of proficiency* at the end of initial training rather than the *type* of initial training used in this study.

Comparison of retraining[4] *procedures.*—Each retention group was split into two subgroups of 5 *S*s each.[5] Assignment to a particular group was on the basis of stanine score at the completion of initial training. This was used as a means of making the subgroups approximately comparable in tracking ability. One group was retrained during four 21-min. sessions with 10 min. rest between sessions; the other during four 21-min. sessions scheduled 1 day apart. Figure 7 presents average performance curves for the retention groups retrained under these two conditions. The varying periods of no practice are equated for the two curves. The two retraining procedures do not result in substantially different performances through the third retraining session, but in the fourth retraining session the distributed practice group appears decidedly superior to the massed practice group.

Apparently, there may be some critical period beyond which performance under massed practice begins to deteriorate. As a means of further evaluating this, the fourth (terminal) session scores for the two retraining procedures were compared statis-

[4] To be more accurate, technically, the term "refresher practice" might be used here since no formal training was involved in these practice sessions. For ease of discussion, however, "retraining" is used.

[5] Since certain cases later were dropped in order to obtain matched retention samples, the subgroups which were used in the data analyses ranged in size from 3 to 5 *S*s.

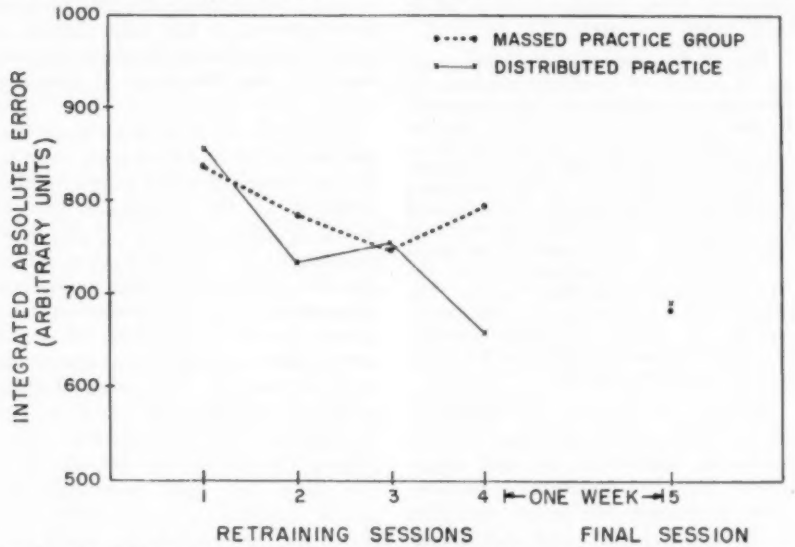


FIG. 7. Effect of different retraining programs during retraining and after a further 1-wk. rest.

TABLE 4

GROUP MEANS AND *SD*s OF INTEGRATED ABSOLUTE ERROR SCORES DURING FINAL TRAINING SESSION, RETRAINING SESSIONS, AND FINAL RETRAINING SESSION

Session	Massed Practice (N = 30)		Distributed Practice (N = 32)	
	Mean	*SD*	Mean	*SD*
Last initial training	733.4	241.9	773.4	229.0
First retraining	836.0	360.6	858.5	308.6
Fourth retraining	801.4	365.4	659.2	232.8
Final (1 wk. later)	684.5	251.0	693.8	225.0

tically. Due to the limited number of cases in each subgroup and the consequent difficulty of matching such scores, an analysis of covariance was conducted which compared fourth session scores while removing the effect of first session scores as a source of variance. In effect, this procedure equates the subgroups statistically and increases the efficiency of the comparison procedure. The results verify the superiority of distributed practice over massed practice as a retraining procedure at the end of four retraining sessions ($F = 10.75$, $df = 1/59$, $P < .01$).

One of the assumptions underlying the use of analysis of covariance concerns the homogeneity of within group regression effect. In this instance it was determined that the two regressions were not homogeneous. This was due to a decrease in the variability of fourth session scores of the distributed practice group as opposed to the massed practice group. To compensate for this, the distributions were transformed by means of a logarithmic transformation and the tests conducted using these scores. The results indicate that the effect of distributed practice, as opposed to massed practice, is both to improve tracking proficiency and to reduce inter-*S* variability.

Relative permanence of retraining benefit.—One week following the final retraining session all *S*s were again tested for one additional session. Table 4 presents the means and *SD*s for both the massed and distributed groups for the first and the final retraining sessions as well as for the session held 1 wk. later. Whereas group performances initially are approximately equal, at the conclusion of the retraining period the distributed practice group is considerably more proficient. However, 1 week later the two groups again are performing at an approximately equal level and this level is closer to that attained by the distributed practice group at the end of retraining than it is to that of the massed practice group. These differences were evaluated statistically; again, an analysis of covariance procedure was used comparing the two groups for the final session with a control on individual variation during the first retraining session. An F of .005 ($df = 1/59$) clearly indicated that during the final session there was no significant difference between the groups. It appears that the differences observed during retraining are not due to differential *learning*, but to temporary factors affecting performance. Thus the same "massed practice, postrest recovery" phenomena are found to occur in *relearning* as has been found repeatedly in studies of original motor learning (see Bilodeau & Bilodeau, 1961).

It is also interesting to note that the performance of both groups during the later session was superior to that attained at the conclusion of the initial training period. Tests were conducted to evaluate this effect. Results indicated a significant improvement for the distributed practice group ($t = 2.57$, $df = 31$, $P = .02$) and similar though not significant improvement for the massed practice group ($t = 1.56$, $df = 29$, $P = .12$). Apparently the five sessions comprising the retraining and the later test trial not only recovered the initial performance capability for these *S*s, but produced additional improvement.

Predicting retention from ability measures.—Those *Ss* in Group I had been administered a battery of 44 printed and psychomotor aptitude tests in an earlier study (Parker & Fleishman, 1959, 1960). A subsequent factor analysis of the correlations among these tests identified 15 ability factors, but only 2 of these (Spatial Orientation and Multilimb Coordination) were found related to performance on the tracking task during initial learning. And these two factors, jointly, never contributed more than 25% of the variance in performance at any stage of practice with this task. Nevertheless, it was thought useful to see if measures of these factors were related to performance after periods of *no* practice.

From their loadings on the two factors (see Parker & Fleishman, 1960), the Stick and Rudder Orientation (printed) Test and the Rudder Control (apparatus) Test were chosen to represent the Spatial Orientation and Multilimb Coordination factors, respectively. Correlations between these tests and performance during the first retention session were computed based on an N of 69 (the Group I *Ss* represented in Table 3, plus the 40 *Ss* brought back for a single session of retention testing). These zero-order correlations with retention performance were .21 for the Spatial test and .18 for the Coordination test; these coefficients are significant at the .10 but not the .05 level of confidence. To hold the effects of initial learning level and retention interval constant, second-order partial correlations were computed. With these factors partialed out the Spatial test correlated .21 and the Multilimb Coordination test correlated .20 with performance in the retention session. Again these coefficients are significant only at the .10 level, not at the .05 level, for second-order partials.

Thus, for this particular skill, a negligible to insignificant portion of retention performance is attributable to *Ss*' abilities as measured prior to initial learning.[6] This is true when retention is defined in terms of performance after no practice, as well as when this performance is residualized with respect to initial learning level and retention interval.

Performance on this task during early stages of *initial* learning was shown to be uncorrelated with performance during late stages of initial learning (e.g., as late as Trial 8 the correlation with Trial 50 was only .13); however, practice sessions late in original learning correlated .70 with each other (Parker & Fleishman, 1960). The communality of the final initial learning trial attributable to independently measured ability factors was only .24. Taken together, these findings suggested that proficiency at the end of training was mainly a function of specific habits and skills acquired during the 6 wk. of practice with the task and only to a small extent a function of *Ss*' abilities prior to his experience with this task.

The present findings indicate this is also true of *retention* performance after prolonged periods of no practice. This is especially apparent when we recall the high correlations (in the .80s and .90s) between proficiency at the conclusion of training and retention performance, relative to the negligible correlations of retention with the independent ability measures.

Summary

Two groups of *Ss* were given extended training on a highly complex tracking task. Practice extended over 17 sessions distributed over 6 weeks. The two groups differed only in the amount of verbal guidance provided in initial training. Within each group, subgroups of *Ss* matched for final proficiency were retested following various no-practice intervals of up to 24 mo. These retention samples were further divided into two subgroups, each of which were given four additional retraining sessions; in one group this relearning practice was massed in 1 day and

[6] The distinction between the constructs "ability" and "skill" has been elaborated elsewhere (Fleishman, 1959, 1962; Gagné & Fleishman, 1959). Briefly, "ability" refers to a more general, stable trait of the individual inferred from response consistencies on a given range of tasks. Skill refers to proficiency on a specific task. Some portion of the variance in a given skill can be accounted for in terms of particular component abilities.

for the other group it was distributed over 4 days. One week following the retraining all *Ss* were retested as a means of evaluating the persistence of the effects of these two relearning schedules.

1. The retention of proficiency in a complex, continuous control, perceptual-motor skill is extremely high, even for no-practice intervals up to 24 mo. For *Ss* trained initially to high levels of proficiency (Group II), virtually no loss was observed for periods up to 14 mo. What small losses did occur were recovered in the first few minutes of relearning. With 24 mo. of no practice, rapid recovery still occurred during the first 20 min. of relearning.

2. Variations in retention interval from 1 to 14 mo. are shown to be unrelated to retention performance, even during the first 1 min. of relearning. The function has zero slope until the loss in performance shown by the 24-mo. retention group.

3. The most important factor in retention is the *level* of proficiency achieved by the *Ss* during initial learning. This effect is shown to be just as important following long and short periods of no practice.

4. The *type* of initial training (amount of verbal guidance) is unrelated to retention performance when proficiency level after original learning is held constant.

5. Retraining administered under conditions of distributed practice proved to be superior to that administered under mass practice based upon a measure of performance during the final retraining session. However, on retesting 1 week later no difference was noted between the two retraining procedures. Thus, in terms of *transfer* to later performance there was no "permanent" disadvantage in massed retraining. Furthermore, both groups had improved beyond their original learning levels.

6. Predictions of individual differences in retention from independent ability measures were negligible. Retention appears more a function of specific task habits acquired, than of *Ss'* ability traits developed prior to training.

References

Ammons, R. B., Farr, R. G., Bloch, E., Neumann, E., Dey, M., Marion, R., & Ammons, C. H. Long-term retention of perceptual-motor skills. *J. exp. Psychol.*, 1958, **55**, 318–328.

Battig, W. F., Nagel, E. H., Voss, J. F., & Brogden, W. J. Transfer and retention of bidimensional compensatory tracking after extended practice. *Amer. J. Psychol.*, 1957, **70**, 75–80.

Bell, H. M. Retention of pursuit rotor skill after one year. *J. exp. Psychol.*, 1950, **40**, 648–649.

Bilodeau, E. A., & Bilodeau, I. McD. Motor skill learning. *Annu. Rev. Psychol.*, 1961, **12**, 243–280.

Fleishman, E. A. Abilities and the learning of psychomotor skills. In P. H. Dubois, W. H. Manning, and C. J. Spies (Eds.), *Factor analysis and related techniques in the study of learning*. St. Louis: Washington University, 1959.

Fleishman, E. A. The description and prediction of perceptual-motor skill learning. In R. Glaser (Ed.), *Training research and education*. Pittsburgh: Univer. Pittsburgh Press, 1962.

Gagné, R. M., & Fleishman, E. A. *Psychology and human performance*. New York: Holt, 1959.

Jahnke, J. C. Retention in motor learning as a function of amount of practice and rest. *J. exp. Psychol.*, 1958, **55**, 270-273.

Jones, E. I., & Bilodeau, E. A. Retention and relearning of a complex perceptual-motor skill after ten months of no practice. *HumRRO res. Bull.*, 1953, No. 53-15.

Leavitt, H. J., & Schlosberg, H. The retention of verbal and of motor skills. *J. exp. Psychol.*, 1944, **34**, 404–417.

Mengelkoch, R. F., Adams, J. A., & Gainer, C. A. The forgetting of instrument flying skills as a function of the level of initial proficiency. *USN Train. Dev. Cent. tech. Rep.*, 1958, No. 71-16-18.

Parker, J. F., & Fleishman, E. A. Prediction of advanced levels of proficiency in a complex tracking task. *USAF WADC tech. Rep.*, 1959, No. 59-255.

Parker, J. F., & Fleishman, E. A. Ability factors and component performance measures as predictors of complex tracking behavior. *Psychol. Monogr.*, 1960, **74**(16, Whole No. 503).

Parker, J. F., & Fleishman, E. A. Use of analytical information concerning task requirements to increase the effectiveness of skill training. *J. appl. Psychol.*, 1961, **45**, 295–302.

Reynolds, B., & Bilodeau, I. McD. Acquisition and retention of three psychomotor tests as a function of distribution of practice during acquisition. *J. exp. Psychol.*, 1952, **44**, 19–26.

(Early publication received February 8, 1962)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 227-233

DISCRIMINATION OF THE REWARD IN LEARNING WITH PARTIAL AND CONTINUOUS REINFORCEMENT [1]

STEWART H. HULSE

Johns Hopkins University

Over the past 20 years, psychologists have spent much time observing and manipulating the things that happen on nonreinforced trials in partial reinforcement learning situations. The result has been a massive array of information, and an equally massive array of theory, concerning the phenomena that partial reinforcement produces (Jenkins & Stanley, 1950; Lewis, 1960). The nonreinforced trial is the hallmark of partially reinforced learning, and the study of the stimulus and response correlates of nonreinforcement is, consequently, of unquestioned importance. But the nonreinforced trial is not the only place to look for information about the effects partial reinforcement produces.

It is quite conceivable that new information could be gained about partial reinforcement from a study of the behavior used to ingest the reward on reinforced trials. This notion is particularly appealing, first of all, in view of recent studies which have shown that consummatory responding, i.e., the rat's licking rate, varies quite systematically as a function of certain variables such as sweetness of reward, size of drop delivered from the drinking tube, and so on (Hulse & Bacon, 1962; Hulse, Snyder, & Bacon, 1960).

Consummatory behavior has another unique property in partial reinforcement situations which makes its study potentially important. Practice of the learned response, running in an alley for example, is unavoidably confounded with number of reinforced and nonreinforced trials, but practice of the consummatory response in the goal box is not. This is true since the consummatory response cannot occur in overt and nonfractional form on nonreinforced trials. The reward is not there. Generally speaking, then, consummatory behavior takes place only as a function of number of reinforced trials. It follows that, while an examination of the development of the running response as a function of number of reinforcements is not logically justifiable, a similar examination of consummatory behavior is justifiable and might be quite interesting.

While the above is true in theory, in fact, some experiments (e.g., Marx, 1958) permit confounding of consummatory behavior and running in the sense that their procedures call for leaving a food cup, a drinking tube, or some other stimulus closely tied to the consummatory response, in the goal box on nonreinforced trials. Presumably, the rat makes abortive licks at the drinking tube, bites the food cup, or otherwise shows fractional components of the consummatory response. Other experimental procedures (e.g., Hulse, 1958; Weinstock, 1954) do, indeed, completely unconfound consummatory behavior from practice of the running response by deliberately removing

[1] This research was supported by National Science Foundation Research Grants G8712 and G18125. The author wishes to thank W. E. Bacon and H. L. Snyder for collecting the data.

the food cup or drinking tube on nonreinforced trials. In both cases, the rationale is generally that of controlling secondary reinforcement or some other process presumed to transpire in the goal box. In neither case has an attempt been made to examine the specific behavior which the goal-box stimuli elicit on reinforced and nonreinforced trials.

The purpose of the present experiment was to examine consummatory behavior on both reinforced and nonreinforced trials as a function of three percentages of reinforcement of a running response in an alley. The confounding of consummatory behavior with practice of running in the alley was deliberately manipulated by making a drinking tube either available to *S* or unavailable to *S* on nonreinforced trials.

Method

Subjects and Apparatus

The *S*s were 50 male naive albino rats of the Sprague-Dawley strain obtained from Sprague-Dawley, Incorporated, Madison, Wisconsin. The *S*s were 70 to 80 days of age at the time they started the experiment.

The apparatus was a 15-ft. U shaped enclosed runway. The sides of the U were 6 ft. long, and the base of the U was 3 ft. long. The start box was 7 in. long and 5½ in. wide, and the goal box was 20 in. long and 4 in. wide. The goal box was attached at right angles to the end of the runway such that *S* made a left-hand turn into it. Inside height of the runway was 4 in. throughout. The runway was covered with hinged pieces of Plexiglas. The goal-box floor was covered with brass shim stock. The apparatus was painted flat black throughout.

Two guillotine doors separated the start and goal boxes from the alley. The goal-box door was located 4 in. before the turn into the goal box proper. Photocells, used to facilitate the timing of running behavior, were located 2 in. past each door.

Reinforcements were provided from a brass drinking tube (2 mm. inside diameter) located behind a Plexiglas shield at the far wall of the goal box. The tube was centered in a ¾ × ⅜ in. vertical slot cut into the Plexiglas. A piece of Masonite could be placed over the entire far wall of the goal box such that *S* could not see or reach the drinking tube on a particular trial.

Each lick on the drinking tube operated a pump system which delivered a drop of water of specified volume to the tip of the tube. The pump system, described in detail elsewhere (Hulse, 1960), consisted of an infusion pump operated by an electronic relay (Otis & Boenning, 1959). The electronic relay also operated a counter.

Procedure

Experimental design.—Three percentages of reinforcement of the running response were used: 33%, 66%, and 100%. In addition, for the partial groups, the drinking tube was available to *S* on nonreinforced trials (T) or it was not available to *S* on nonreinforced trials (NT). An NT condition could not be included for the 100% *S*s, since this percentage of reinforcement required the presence of the drinking tube on all trials. Ten *S*s were used in each of the five groups called for by the design.

Taming.—On each of 10 taming days, the *S*s were handled freely and placed in groups of 5 or 6 into a large wooden box. Six water bottles were clipped to the outside of the box with their drinking tubes projecting through holes in the walls of the box. The *S*s were permitted to explore and to drink from the tubes for 15 min. This was the only water available during taming. Purina lab chow pellets were available in the individual home cages at all times.

Training.—Following taming, *S*s were given 60 training trials, 1 trial per day. On each trial, *S* was placed in the start box, and after a 2- to 3-sec. delay, the start-box door was raised. The start-box and goal-box doors were lowered after *S* had passed under them. After *S* was removed from the goal box, it was returned to its home cage. Fifteen to 30 min. later, a water bottle was attached to the cage, and *S* drank for 30 min. The *S*s were thus approximately 23 hr. thirsty at the time trials began each day.

All *S*s were reinforced on each of the first 3 days of training. Reinforcement consisted of 600 licks on the drinking tube with the pump set to deliver .0053 cc of water with each lick. On Day 1, the drinking tube projected through the slot into the goal box. On Days 2 and 3, the tube was gradually withdrawn so that on Day 3, its tip was ⅛ in. behind the slot. The tube remained in this position for the rest of the experiment.

On Day 4, the first nonreinforced trial was introduced for the partial *Ss*. Thereafter, reinforced and nonreinforced trials were determined randomly according to the percentages of reinforcement called for by the experimental design. The only other restriction on randomization of reinforcement was that the last training trial was reinforced for all *Ss*.

On nonreinforced trials for the partial *Ss* tested under the T condition, the drinking tube was present in its usual location behind the slot, but the pump system was emptied of water. The *S* was thus free to lick from the tube, but no water was obtainable. On nonreinforced trials for *Ss* tested under the NT condition, the far wall of the goal box was covered with a piece of Masonite so that *S* could neither see nor lick from the tube. Goal-box confinement on nonreinforced trials was for 3 min.

Measures of performance.—Three measures of running behavior were recorded: start time, alley time, and goal-box time. Start time began when the start-box door went up and ended when *S* passed the photocell outside the start-box door. Alley time was the time *S* required to run from the first photocell to the second photocell at the goal-box door. Goal-box time was the time *S* required to run from the second photocell to the drinking tube and make the first lick. Goal-box times were not recorded on nonreinforced trials for the NT groups, since the drinking tube was not available. All times were recorded automatically on Standard Electric timers. The time scores were transformed to reciprocals and multiplied by 100 for purposes of the statistical analyses.

Licking times on reinforced trials were determined from a clock which started with the first lick on the drinking tube and stopped when *S* had completed its 600-lick allotment. These times were transformed to rates of licking, in licks per second. The number of licks *Ss* in the T groups emitted on nonreinforced trials was also recorded.

Results

Licking.—Figure 1 shows that if licking rates are plotted as a function of number of reinforcements, the rates for different percentages of reinforcement reach approximately the same asymptote at the last five reinforced trials. However, the figure also shows that the rate of increase

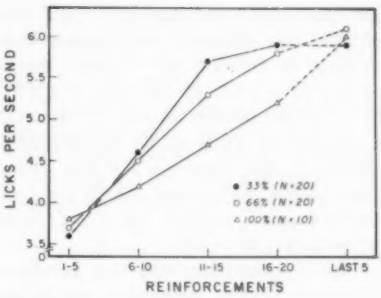


FIG. 1. The development of licking rates as a function of ordinal number of reinforced trials. (The scores have been combined for the partial groups, since the statistical analyses indicated the groups did not differ according to the T and NT variable.)

of licking over the first 20 reinforcements is quite different depending upon percentage of reinforcement. The curve for the 33% groups appears negatively accelerated, that for the 66% groups almost linear with a slight suggestion of negative acceleration, and that for the 100% group positively accelerated.

There is abundant statistical evidence to support the significance of the differences shown in Fig. 1. A simple F test for percentage of reinforcement based on means for the last 10 of the first 20 trials shown in the figure yields an F of 7.92 ($df = 2/47$, $P < .01$). An analysis of variance of 5-trial means for the first 20 reinforcements for the 33% and 66% T groups and the 100% group, i.e., for those groups which had comparable goal-box conditions on all training trials, yields an F for percentage of 4.44 ($df = 2/27$, $P < .05$). This analysis also shows a significant Percentage × Blocks of Trials interaction ($F = 2.50$, $df = 6/81$, $P < .05$) which indicates that the three groups are progressing towards their final asymptote at different rates. A similar analysis based on 5-trial means

for the 33% and 66% T and NT groups also shows a significant Percentage × Blocks of Trials interaction ($F=3.12$, $df=3/108$, $P<.05$). No analysis shows means on the last 5 reinforced trials to be significantly different; the analysis for the combined 33%, 66%, and 100% groups, for example, yielded an F of 3.00 ($df = 2/47$, $P > .05$).

The T condition did not produce different licking behavior on reinforced trials than the NT condition. The analysis of variance of 5-trial means for the 33% and 66% T and NT groups provides no evidence that licking rates differ as a function of the availability of the tube in the goal box on nonreinforced trials ($F=1.66$, $df=1/36$, $P>.05$). This fact is equally true for the last 5 reinforced trials of the 60 training trials ($P > .05$). Moreover, the T and NT condition does not interact with percentage of reinforcement ($P > .05$).

Figure 2 shows that the 66% T group emits more licks on nonreinforced trials than the 33% T group ($F = 6.02$, $df = 1/18$, $P < .05$), number of licks decreases for both groups across blocks of nonreinforced trials ($F = 10.92$, $df = 3/57$, $P < .01$), but

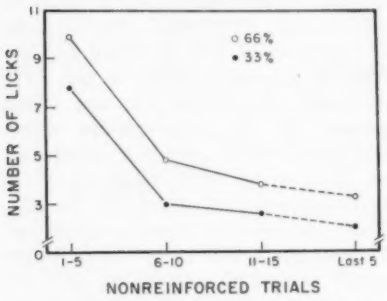


FIG. 2. Number of licks on nonreinforced trials for the 33% and 66% T groups as a function of ordinal number of nonreinforced trials.

the difference in number of licks between the groups across blocks of trials does not change in size ($P>.05$). Of particular importance is the fact that the absolute number of licks for both groups is small, and most of the decrease in number of licks occurs during the course of the first 5 to 10 nonreinforced trials.

Many of the *S*s in the T groups failed to lick at all on some nonreinforced trials. This happened more frequently for the 33% T group than for the 66% T group where, out of the 200 trials represented by the first 15 and last 5 nonreinforced trials, the frequencies are 73 and 60 nonlick trials, respectively. A Wilcoxon T test shows this difference to be significant ($P < .05$).

Running.—We noted that separate analysis of running behavior for reinforced and nonreinforced trials is not logically justifiable, since practice of running is confounded with the number of such trials. Strictly speaking, however, this is not true for goal-box speeds, since this measure reflects behavior which occurs after *S*s enter the goal box and expose themselves directly to the different stimulus conditions correlated with reinforcement and nonreinforcement.

Although an analysis of variance of blocks of 5 trials for the first 20 reinforced trials shows a significant Blocks × Percentage interaction ($F = 2.64$, $df = 6/81$, $P < .05$), goal-box speeds differ only for the first 5 reinforced trials. For this block of trials, speeds are greater the higher the percentage of reinforcement. The groups do not differ on the last 5 reinforced trials of the 60 training trials ($P>.05$).

An analysis of variance of goal-box speeds on the first five and last five nonreinforced trials for the 33% and 66% T groups shows that the 66% group ran to the tube faster than

the 33% group ($F = 7.47$, $df = 1/18$, $P < .05$), but that goal-box speeds for the two groups *decrease* from the beginning to the end of training ($F = 6.70$, $df = 1/18$, $P < .05$). The interaction between the variables is not significant.

The differences obtained for behavior in the goal box are indeed unique to the goal box and did not result from some extraneous factor such as differential handling by *E* at the beginning of reinforced as compared with nonreinforced trials. This is substantiated by a control analysis run on alley speeds for Reinforced Trials 16 to 20 which showed a significant effect for percentage of reinforcement ($F=4.78$,$df=2/47$, $P<.05$). Partially reinforced *S*s run faster than continuously reinforced *S*s, as we would expect, since the former have had many more trials in the alley.

An analysis of start and alley speeds at the end of 60 training trials shows little except that a 15-ft. **U** shaped alley produces a great deal of response variability. Analyses of variance based on means for the last 20 training trials for start speeds show no significant differences due to any of the variables. The same is true for alley speeds, except for an analysis which compared means for the 33% and 66% condition and the T and NT condition. Here, a significant percentage effect was obtained ($F=4.27$, $df = 1/36$, $P < .05$) which indicated that 66% reinforcement produced higher speeds than 33% reinforcement. This effect was not significant, however, for an analysis based on means from the 100%, 66% T, and 33% T groups.

Discussion

Rats quickly learn to detect whether or not the drinking tube will produce water, and their licking rates rapidly increase as a function of number of reinforcements. However, the *rate* with which licking rates increase over reinforced trials is a function of the number of intervening nonreinforced trials.

We can account for this phenomenon on the assumption that a stimulus-discrimination process takes place in the goal box for the partial groups. If we think of the drinking tube and fluid as discriminative stimuli for approaching the tube and drinking, the data suggest that rats quickly learn to attend to these stimuli and to do the appropriate thing when they are present or absent.

First, licking on nonreinforced trials occurs, but it rapidly extinguishes. The *S*s in the T condition learned to take only a lick or two on the dry drinking tube on nonreinforced trials, and they learned to do this early in training over the first 5 to 10 nonreinforced trials. Also, after a given number of nonreinforced trials, the *S*s in the 66% T group have had more reinforced trials than *S*s in the 33% T group and have, presumably, developed a stronger consummatory response. We might, therefore, expect a greater tendency for them to generalize licking on reinforced trials to licking on nonreinforced trials. This apparently occurred, since the 66% T group emitted more licks on nonreinforced trials than the 33% T group. Finally, in this connection, *S*s in the partial groups were apparently so set to discriminate cues associated with reinforcement from cues associated with nonreinforcement that they were able to outwit *E* in their ability to identify a nonfunctional drinking tube. Under the T condition, all *S*s failed to lick on some nonreinforced trials. Possibly, since the pump system was drained on nonreinforced trials, the *S*s could discriminate after they entered the goal box whether or not water was visible at the opening of the tube. Possibly, for the same reason, they discriminated an odor difference in the goal box; given the chemical content of local laboratory tap water, this is not an inconceivable proposition.

Second, goal-box speeds on nonreinforced trials decrease as a function of the number of such trials. Since the licking data show that partial *Ss* could sometimes detect a dry drinking tube before they took a lick, approaching the drinking tube on a nonreinforced trial extinguished to some extent. Moreover, the effects of this process appear to have generalized to suppress the speed with which partial *Ss* approached the drinking tube on reinforced trials. Thus, goal-box speeds on reinforced trials increase at approximately the same rate regardless of percentage of reinforcement. This is to be contrasted with the usual finding, obtained for alley speeds in the present experiment, that increments in response strength as a function of number of reinforced trials will be much greater the lower the percentage of reinforcement of the response.

Finally, the consummatory response develops over reinforced trials as *Ss* learn to discriminate the special significance of the drinking tube and fluid as stimuli for consummatory responding. For a fixed number of reinforcements, the rate with which this discrimination develops will increase as the number of intervening nonreinforcements increases. We would therefore expect the 33% groups to learn the consummatory response faster than the 66% groups. Since the 100% group received no discrimination training with respect to the drinking tube, we would expect this group to develop the consummatory response slowest of all. The data clearly support these conclusions.

It seems clear that partial reinforcement provides discrimination training for reward stimuli in the goal box, but continuous reinforcement does not. During partial reinforcement, in effect, behavior is critically focused on the reward and its stimulus properties because of the contrast in goal-box conditions on reinforced as compared with nonreinforced trials. It follows that, if partial reinforcement is used to condition a response, the development of response strength may be more critically determined by stimuli correlated with the reward, such as its sweetness, than if continuous reinforcement is used. We might expect, for example, that after partial reinforcement resistance to extinction would increase as a function of the sweetness of a reward. After continuous reinforcement, however, there should be much less correlation between reward sweetness and resistance to extinction. Hulse and Bacon (1962) showed these predictions to hold following training with different concentrations of saccharin. Hulse (1958) showed much the same thing for different-sized food rewards.

Some additional support for these hypotheses comes from another source. Jenkins (1961) found that resistance to extinction in the presence of a stimulus was greater if that stimulus had been paired with reinforcement during discrimination training as compared with continuous reinforcement training. Jenkins' stimulus was a light pattern projected on *S*'s response key. The analogous stimulus in the present approach is the reinforcing stimulus itself, and this is not presented, of course, during extinction. If the effect Jenkins noted is to occur in straightforward fashion, it would have to operate through some mechanism other than the reward stimulus per se. There is a clear parallel between Jenkins' approach and the approach outlined here, and the correspondence between Jenkins' data and the data discussed here is interesting. This apparent correspondence must remain suggestive for the present, however.

Summary

Fifty albino rats were given 60 training trials in a 15-ft. U shaped alley. They were reinforced with 600 licks of water from a drinking tube on 33%, 66%, or 100% of the trials. Half the partial *Ss* could lick on the dry drinking tube on nonreinforced trials; for the other half, the tube was blocked such that the *Ss* could neither see it nor lick it. Start speeds, alley speeds, and goal-box speeds, licking rates on reinforced trials, and number of licks on nonreinforced trials were recorded.

The results show that the partial groups

developed licking rates faster in the goal box, as a function of number of reinforcements, than the continuous group. This happened whether or not the tube was available on nonreinforced trials. Licking rapidly extinguished, and goal-box speeds decreased, on nonreinforced trials for the partial groups. Start speeds and alley speeds did not vary as a function of any of the experimental variables; this may have been due to excessive response variability produced by the very long runway.

The data suggest that partial reinforcement produces a very powerful discrimination of reward stimuli in the goal box. This process may be a factor in experiments which have shown that stimulus variables correlated with the reward, such as its sweetness or size, have different effects if they are used with partial as compared with continuous reinforcement.

References

Hulse, S. H. Amount and percentage of reinforcement and duration of goal confinement in conditioning and extinction. *J. exp. Psychol.*, 1958, **56**, 48–57.

Hulse, S. H. A precision liquid feeding system controlled by licking behavior. *J. exp. Anal. Behav.*, 1960, **3**, 1–3.

Hulse, S. H., & Bacon, W. E. Supplementary report: Partial reinforcement and amount of reinforcement as determinants of instrumental licking rates. *J. exp. Psychol.*, 1962, **63**, 214–215.

Hulse, S. H., Snyder, H. L., & Bacon, W. E. Instrumental licking behavior as a function of schedule, volume, and concentration of a saccharine reinforcer. *J. exp. Psychol.*, 1960, **60**, 359–364.

Jenkins, H. M. The effect of discrimination training on extinction. *J. exp. Psychol.*, 1961, **61**, 111–121.

Jenkins, W. O., & Stanley, J. C., Jr. Partial reinforcement: A review and critique. *Psychol. Bull.*, 1950, **47**, 193–234.

Lewis, D. J. Partial reinforcement: A selective review of the literature since 1950. *Psychol. Bull.*, 1960, **57**, 1–28.

Marx, M. H. Resistance to extinction as a function of continuous or intermittent presentation of a training cue. *J. exp. Psychol.*, 1958, **56**, 251–255.

Otis, L. S., & Boenning, R. A. A transistorized circuit for recording contact responses. *J. exp. Anal. Behav.*, 1959, **2**, 289–291.

Weinstock, S. Resistance to extinction of a running response following partial reinforcement under widely spaced trials. *J. comp. physiol. Psychol.*, 1954, **47**, 318–322.

(Early publication received March 16, 1962)

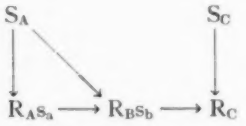
Journal of Experimental Psychology
1962, Vol. 64, No. 3, 234–238

MEDIATED ASSOCIATION IN A PAIRED-ASSOCIATE TRANSFER TASK

DAVID S. PALERMO

Institute of Child Development, University of Minnesota

A reliable condition for producing negative transfer in paired-associate learning occurs in the A-B, A-C paradigm. In this paradigm, one set of responses is learned to a set of stimuli and subsequently, new responses are learned to the same stimuli. Another approach to an understanding of the learning of successive lists of paired associates has used an A-B, B-C, A-C paradigm for an experimental group and an A-B, D-C, A-C paradigm for a control group. In such mediated association studies (e.g., Norcross & Spiker, 1958), it has been assumed that the A-C list is learned more rapidly by the experimental group because the learning of the B-C list provides a mediating link between A and C of the following nature:



The $s_b \rightarrow R_C$ association learned in the B-C list provides the associative or mediational link assumed to facilitate the learning of the A-C list.

The present experiment was designed to determine whether mediated association effects can be demonstrated within an A-B, A-C paradigm. The usual paradigm was modified to the extent of presenting lists in which each response was paired with two stimuli rather than one response with each stimulus. Therefore, there were only half as many different responses as stimuli in both the A-B and the A-C lists. Upon reaching criterion on the A-B list, each response and its associated stimulation had been conditioned to two stimuli of the list. By appropriate pairing of the new responses with the old stimuli in a "mixed" A-C list (Twedt & Underwood, 1959), it was possible to arrange stimuli and responses which, according to the mediational hypothesis, would lead to facilitation or impairment of parts of the A-C list *during* the learning of the list. Thus, the learning of a particular pair in the A-C list was assumed to mediate responses which would affect the learning of another pair within the list.

METHOD

Apparatus.—The apparatus was a modified Card Master. Basically, it consisted of a box 13 × 13 × 9½ in. The front face of the apparatus contained a 3 × 6 in. aperture covered by clear plastic. A mechanical arrangement delivered to the aperture a 3½ × 6 in. plastic card from the bottom of a stack of cards. Mounted on each card was a pair of Stanford-Binet Picture Vocabulary pictures reproduced by a Thermofax process for this purpose. Gray metal doors served as shutters to expose independently the left and right pictures on the cards. Following a presentation, the doors closed simultaneously and the card was released from the aperture and returned to the top of the stack by a conveyor belt. A system of electronic timers controlled the rate of presenting the cards and the exposure times of the stimulus and response picture.

Experimental design.—All *S*s learned two lists, each of which was composed of six stimulus pictures and three response pictures paired so that each response was learned to two stimuli. Table 1 presents the design of the experiment, with the picture names, which provides that each *S* serve under each

TABLE 1

DESIGN OF THE EXPERIMENT

List I		List II		Cond.
S	R	S	R	
S_1 STOOL	R_A SCISSORS	S_1 STOOL	R_D HOUSE	I
S_2 SHOE	R_A SCISSORS	S_2 SHOE	R_D HOUSE	I
S_3 GLASSES	R_B BED	S_3 GLASSES	R_E BASKET	II
S_4 CUP	R_B BED	S_4 CUP	R_F KNIFE	II
S_5 CLOCK	R_C HAND	S_7 FORK	R_E BASKET	III
S_6 TABLE	R_C HAND	S_8 TREE	R_F KNIFE	III

of three experimental conditions. For two pairs in List 2 the mediational chain, established during the learning of the pairs, is expected to facilitate learning (Cond. I). For two other pairs the mediational link is expected to elicit incorrect responses and, thus, interfere with learning (Cond. II). The other two pairs involve two new stimuli as well as the two new responses and thus have no experimentally established mediating tendencies (Cond. III). In the case of Cond. I and II, it is assumed that when a stimulus is presented in List 2, the List 1 response and attendant stimulation will occur either overtly or covertly. It is also assumed that the response of List 2 will be conditioned to the attendant stimulation of the List 1 response.

For example, in List 2 of Table 1 the learning of the pair STOOL-HOUSE ($S_1 \rightarrow R_D$) in Cond. I is expected, through the establishment of the association SCISSORS-HOUSE ($s_a \rightarrow R_D$), to increase the tendency to make the response HOUSE (R_D) to SHOE (S_2). Similarly, the learning of SHOE-HOUSE ($S_2 \rightarrow R_D$) is expected to facilitate the learning of STOOL-HOUSE ($S_1 \rightarrow R_D$). In the case of Cond. II, it is expected that the learning of GLASSES-BASKET ($S_3 \rightarrow R_E$), in List 2, will interfere with the learning of CUP-KNIFE ($S_4 \rightarrow R_F$) because of the establishment of the association of bed-BASKET ($s_b \rightarrow R_E$). Similarly, the establishment of the association bed-KNIFE ($s_b \rightarrow R_F$) during the learning of CUP-KNIFE ($S_4 \rightarrow R_F$) will interfere with the learning of GLASSES-BASKET ($S_3 \rightarrow R_E$). The occurrence of BASKET (R_E) to CUP (S_4) or KNIFE (R_F) to GLASSES (S_3) would be considered a mediated error. In the case of Cond. III, the stimuli in List 2 have no experimentally established mediating tendencies since they have not been previously presented.

Subjects.—The *Ss* were drawn from Grades 3–6 (*Ns* = 12, 12, 15, and 18, respectively) of the University of Minnesota Elementary School.[1] All *Ss* were naive with respect to verbal learning except the Grade 6 children who had participated in a paired-associate learning experiment 1 yr. previously. Seven *Ss* had to be discarded due to apparatus failure and 19 were discarded because they failed to reach criterion on List 1 within 27 trials (11, 4, 4, and 1 in Grades 3–6, respectively). All groups received the same form of List 1. One-third of the *Ss* in each grade level were randomly assigned to each of three groups designed to control for possible differential difficulty of the stimulus-response pairs in List 2. The groups differed with respect to which of the three forms of List 2 they received. The three forms were constructed so that a given set of two List 2 pairs was used in each of the three conditions.

Procedure.—The *Ss* were given two experimental sessions separated by 1–4 days. Each *S* was brought by *E* from the classroom to the experimental room, and seated facing a 37 × 49 in. panel with a 3 × 6 in. aperture through which *S* viewed the stimulus cards. A shaded 40-w. light directly above the aperture provided illumination. The apparatus was located behind the panel in an adjoining room. On the table was a Webster Teletalk two-way speaker. The *E* gave the instructions through the speaker system and recorded the verbal responses as they were given by *S*. Thus, *S* was alone during both experimental sessions.

At Session 1 instructions explaining the task were read to *S* and a pair of pictures, not subsequently used in the experiment, were shown to familiarize him with the apparatus

[1] The author wishes to express his appreciation to James R. Curtin, Principal of the University Elementary School, for his cooperation in making *Ss* available for the experiment, and Ernest Washington, who ran the *Ss*.

and procedure. The *S* was not required to name the stimulus picture but was asked to anticipate the response picture by naming it. No *S* had any difficulty in giving the appropriate name to the stimuli. List 1 was presented immediately and continued until *S* reached a criterion of three successive errorless trials.

At Session 2 each *S* relearned List 1 to three successive errorless trials and then was told that another list would be presented. After approximately 2 min., necessary to change lists, List 2 was presented for nine trials or three successive errorless trials, whichever was fewer. If the criterion was reached prior to the nine trials, it was assumed for the analysis that no additional errors would have been made.

The pictures in List 1 were randomly designated as stimuli and responses and randomly paired together with the restriction that no obviously highly associated pictures were paired together. All lists were presented in three varying orders to control for serial learning. The orders were randomly determined with the exception that no response picture was presented twice in succession. A 3-sec. anticipation period, a 3-sec. joint presentation of stimulus and response, and a 3-sec. interval between pairs was used throughout the experiment. All verbal responses occurring in the anticipation interval were recorded verbatim throughout the experiment.

Results

Table 2 presents the means and *SD*s for the number of correct anticipations in each condition for each grade. An evaluation of the differences among these groups employed an analysis of variance in which the main effects of Control List Groups and of Grade were between-*S* factors while the main effect of Conditions was an intra-*S* factor (Lindquist, 1953, p. 281, Type III). The means for the groups designed to control for possible differential difficulty of the picture pairs did not differ significantly ($F = 1.68$, $df = 2/45$, $\text{Error}_b = 14.54$), nor did the four grades ($F = 2.56$, $df = 3/45$, $.05 < P < .10$). The means for the three experimental conditions differed beyond the .001 level ($F = 14.88$, $df = 2/90$, $\text{Error}_w = 3.32$). Individual t tests for related measures showed that the differences between Cond. I and III and between Cond. II and III were significant ($P < .01$), but the difference between Cond. I and II was not ($.05 < P < .10$). None of the interactions was significant.

Although the analysis of variance indicated no significant interaction between Grade and Conditions, it may be seen in Table 2 that there appears to be an interaction between Cond. I and II and grade level. Performance of *S*s in Grades 3 and 4 shows clearly the effects of the two conditions, but for Grade 5 the effect is much less pronounced and for

TABLE 2

Means and *SD*s of Trials to Learn List 1 and for Correct Responses on List 2 under Each Condition for Each Grade

Grade	*N*	List 1		List 2							
		Trials to Learn		Cond. I		Cond. II		Cond. III		Cond. I, II, III	
		Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
3	12	19.42	7.64	11.58	3.10	10.33	3.40	12.50	2.81	11.47	3.24
4	12	19.42	7.27	12.92	2.50	10.83	2.11	14.33	2.28	12.69	2.72
5	15	19.27	6.51	12.61	3.26	12.27	2.52	13.27	2.77	12.71	2.89
6	18	12.39	3.27	13.11	2.08	13.44	2.09	14.67	1.20	13.74	1.95
Total	57	17.16	6.96	12.61	2.80	11.93	2.82	13.77	2.44	12.77	2.80

Grade 6 the means are actually in the opposite direction of that predicted. Consequently, an analysis of variance was conducted on the data for Cond. I and II alone to permit a more sensitive test for the interaction indicated in Table 2. The difference between the conditions was significant at the .05 level ($F = 4.14, df = 1/45, \text{Error}_w = 3.22$), but the Grade × Conditions interaction was not significant by usual standards ($F = 2.48$, $df = 3/45$, $.05 < P < .10$).

An analysis of the errors made in response to stimulus members of the pairs in Cond. II and III was made by the procedure used by Norcross and Spiker (1958). The mean number of reversals per *S* for Cond. II was 1.33 and for Cond. III, .93 ($P = .10$, related t test). For Grades 3 and 4, where the mediation effect was more apparent, there were 1.58 reversals in Cond. II and 1.00 in Cond. III ($P = .05$, related t test). However, these suggestions of differences must be discounted because the ratio of the number of such intrusion errors to total errors, was approximately equal in each comparison of Cond. II and III.

Discussion

These data provide evidence that the conditions designed to produce mediated associations result in differential effects upon the amount of transfer in an A-B, A-C transfer paradigm. Performance on the pairs assumed to be facilitated by the mediation of correct responses was superior to performance on pairs in which interference was assumed due to the mediation of incorrect responses. This result had been found in a preliminary study using 22 Grade 6 *S*s in which List 2 was composed of four pairs designed to produce facilitation through mediation (Cond. I) and four pairs designed to produce interference through mediation (Cond. II) but without the control pairs (Cond. III). The mean number of correct responses for Cond. I and II in that study was 25.00 ($SD = 4.00$) and 23.32 ($SD = 3.76$), respectively, yielding a t which was significant at the .025 level.

While the relationship between the conditions assumed to produce facilitation and interference due to mediation seems to be a stable one under these conditions, some caution should be made in generalizing these results. Attempts to replicate the findings of this study with college students and Grade 10 children using the same design failed to yield significant differences among the three conditions. In the study with college students, low-association nonsense syllables were used as stimuli and responses, and in the Grade 10 study high-frequency adjectives were used.

The suggestion of an interaction between grade level and the effects of Cond. I and II in this study along with the failure to replicate the findings with college students and tenth graders might indicate that some factor or factors related to age may be important.

In the present study, performance on the control pairs was superior to performance on the pairs assumed to be facilitated by mediation as well as on the pairs assumed to be impaired by mediation. These results are in contrast with those of Norcross and Spiker (1958). Performance on their control pairs was superior to performance on pairs in which interference was expected and inferior to performance on pairs in which facilitation was expected in a three-list paradigm. Apparently, the strong associative interference effects of the A-B, A-C paradigm in the present study outweigh any facilitation introduced by the mediated associations.

Ignoring the mediation effects, this study replicates the findings of the Spiker and Holton (1958) study in which it was demonstrated, using a motor paired-associate task, that learning of an A-B, A-C series results in associative interference relative to an A-B, D-C series. In the present experiment the control pairs are comparable to the D-C pairs

of the Spiker and Holton study and were learned significantly faster than either the Cond. I or II pairs.

Summary

An experiment was conducted to study facilitation and impairment of performance as a function of mediated associations in a modified A-B, A-C transfer paradigm. Children in Grades 3, 4, 5, and 6 were required to learn two lists of six paired associates composed of six stimuli and three responses. In List 2, the pairs were arranged so that learning could be facilitated by mediated associations or impaired by mediated associations. In addition, control pairs, in which no experimentally manipulated mediation was present, were included.

The results indicated that the condition designed to produce facilitation through mediated associations led to superior performance when compared with the condition designed to produce impairment through mediated associations. Performance on the control pairs was superior, however, to that of both of the mediation conditions. The results were discussed in terms of age related variables and the associative interference effects of the A-B, A-C paradigm.

References

Lindquist, E. F. *Design and analysis of experiments in psychology and education.* New York: Houghton Mifflin, 1953.

Norcross, K. J., & Spiker, C. C. The effects of mediated associations on transfer in paired-associate learning. *J. exp. Psychol.*, 1958, **55**, 129–134.

Spiker, C. C., & Holton, R. B. Associative transfer in motor paired-associate learning as a function of amount of first-task practice. *J. exp. Psychol.*, 1958, **56**, 123–132.

Twedt, H. M., & Underwood, B. J. Mixed vs. unmixed lists in transfer studies. *J. exp. Psychol.*, 1959, **58**, 111–116.

(Received July 17, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 239–248

PREDICTION OF PREFERENCE, TRANSPOSITION, AND TRANSPOSITION-REVERSAL FROM THE GENERALIZATION GRADIENT [1]

WERNER K. HONIG

Denison University

The present study concerns the prediction of differential responding in an operant situation from the generalization gradient. Gradients were obtained on the spectral continuum with the technique of Guttman and Kalish (1956), while preference tests involving the generalization stimuli were concurrently administered. In a general sense, this is an attempt to predict choice behavior when the responsiveness to each single stimulus is known. More specifically, it involves a test of the power of the generalization gradient in the determination of behavior in situations more complex than the single-stimulus condition in which it is assessed.

The desired relationships were investigated after simultaneous and successive[2] discrimination training as well as simple acquisition, in order to determine whether the peak shift obtained by Hanson (1959) for the postdiscrimination gradient (PDG) is accompanied by transposition and transposition-reversal as predicted by Spence (1937). In previous transposition studies (e.g., Baker & Lawrence, 1951; Ehrenfreund, 1952; Kendler, 1950) the PDGs have not been obtained; it could not be determined, therefore, whether the consistent failure to obtain clear-cut transposition-reversal in such studies was due to the form of the PDGs or to a lack of correspondence between them and stimulus preferences on the transposition tests. In the present study, PDGs and transposition tests were obtained concurrently.

EXPERIMENTS 1 AND 2

Method

Apparatus.—The automatic key pecking apparatus used in the present investigation was similar to that employed in other spectral generalization studies (e.g., Guttman & Kalish, 1956) except that it included two separately illuminated keys 7½ in. from the floor with 2 in. between centers. Bausch and Lomb monochromatic interference filters, with band widths at half height of 7–9 mμ, provided the different spectral values. In the text, stimuli will be referred to by the nominal value of the filters, but the spacing of points on the abscissae of the figures corresponds to the actual transmission peaks which differ slightly in some cases. A K-2 yellow filter was inserted when necessary to prevent the transmission of the visible second-order spectrum produced by the filters. The various spectral values were equated for apparent brightness by a human observer with the aid of a Macbeth illuminometer. This appears justified by the work of

[1] This paper is adapted from a doctoral dissertation submitted to the Department of Psychology, Graduate School of Arts and Sciences, Duke University, in 1958. The research was supported by Grant MH-1002 from the National Institute of Mental Health to Norman Guttman. The author is much indebted to Norman Guttman for advice and guidance throughout the research. The preparation of the publication version was supported by Grant M-2414 to the author.

[2] A *successive discrimination* refers in this paper to the case where only the positive or the negative stimulus value is presented at one time to *S*, or the "go, no-go" discrimination. The term has been used in this sense by Grice (1949) and Baker and Lawrence (1951). Other authors (Bitterman, Spence) have more recently used *successive discrimination* to refer to a conditional left-right discrimination, but this usage is not intended here.

TABLE 1

COMPOSITION OF STIMULUS PAIRS ON THE GENERALIZATION TESTS IN EXP. 1 AND 2

Experiment 1		Experiment 2	
Absolute Value (mμ)	Difference from CS (mμ)	Absolute Value (mμ)	Difference from CS (mμ)
490, 510	−60, −40	490, 590	−60, +40
510, 520	−40, −30	590, 520	+40, −30
520, 530	−30, −20	520, 570	−30, +20
530, 540	−20, −10	570, 540	+20, −10
540, 550	−10, 0	540, 550	−10, 0
550, 560	0, +10	550, 560	0, +10
560, 570	+10, +20	560, 530	+10, −20
570, 580	+20, +30	530, 580	−20, +30
580, 590	+30, +40	580, 510	+30, −40
590, 610	+40, +60	510, 610	−40, +60

Blough (1957), which indicates that the spectral luminosity function of the pigeon is similar to that of man. No illumination was provided in the box other than the light falling on the keys, except during the presentation of grain, when a magazine light went on.

Subjects.—The *S*s were 38 white Carneau pigeons reduced to 75% of their free-feeding weight. Twenty-two experimentally naive animals were used in Exp. 1. The 6 *S*s used in Exp. 2 had served in a previous experiment in which 10 sessions of pecking at a 550-mμ stimulus on a variable interval schedule were followed by a generalization test in which the values used in the present study were presented. The previous experiment was run in a different apparatus which had one key.

Procedure.—Over several daily sessions, *S* was magazine trained, conditioned to peck at a 550-mμ key, and given 60 continuous reinforcements. Following this, *S* received 10 daily sessions of variable interval (VI) training with a mean interreinforcement interval of about 1 min. Each session consisted of 30 1-min. periods separated by 10 sec. of blackout. During half the periods, one key was illuminated by 550 mμ. During the remaining periods, both keys were illuminated by 550 mμ. The single stimulus appeared for half the time on the right and for half the time on the left key. The appearance of the single stimulus, and the availability of reinforcement in the two-stimulus case, were balanced between left and right keys. The order of conditions was randomized. Five of the 22 *S*s in Exp. 1 received training beyond 10 sessions until they reached a criterion of 900 responses in a single session.

This training procedure was abbreviated for Exp. 2, as magazine training and conditioning of key pecking could be omitted. After 10 continuous reinforcements, VI training began immediately in the manner described above. All *S*s received 7 rather than 10 sessions of VI training, and easily reached 900 responses per session by the seventh session.

On the 2 days following the last session of VI training, *S* received a generalization test under extinction, consisting of 189 30-sec. periods of stimulus presentation alternating with 10 sec. of blackout. Six blocks of 21 periods each were given on the first day of testing and three on the second. A single stimulus value was presented on 11 of the 21 periods within each block and a pair of stimulus values on the remaining 10 periods. The single stimulus values ranged from 490 to 610 mμ in 10-mμ steps, with the omission of 500 and 600 mμ. The stimulus pairs were composed of these values as summarized in Table 1. In Exp. 1, the pairs were composed of adjacent stimuli in the series. In Exp. 2, they were so chosen that the differences in mμ between the test values and the CS in each pair correspond to those for Exp. 1 in size, but have opposite signs. In other words, in Exp. 1 the stimuli in each pair lie on the "same side" of the CS, while in Exp. 2 they lie on "opposite sides."

The 21 stimulus conditions were randomized within each block of presentations. Within a given block, 490, 520, 540, 560, 580, and 610 mμ were presented on one key and 510, 530, 550, 570, and 590 mμ were presented on the other key. This arrangement alternated on successive blocks.

By means of this testing procedure, it was possible to present a generalization test consisting of singly presented stimulus values in the same session with a test consisting of different pairs of values. These will be distinguished as single-stimulus (SS) and double-stimulus (DS) generalization tests, even though they occurred in the same session. The testing procedure provided control for differences between animals and for the effects of extinction in the course of testing.

Results

Experiment 1.—The data obtained from the generalization tests are presented in Fig. 1. The mean total responses to SS values are shown by

the single-stimulus gradient (SSG); the mean total responses to the members of each stimulus pair are shown by the adjacent bars. The filled bar in each pair represents the stimulus nearer the CS. When the responses to each stimulus *value* on the DS test are summed across the two pairs in which that value appears, the double-stimulus gradient (DSG) is obtained.

The SSG peaks at 550 mμ and decreases to both sides of the CS. There are inversions between 490 and 510 mμ and between 560 and 570 mμ. The former inversion is small and may be due to random error, as the gradient is almost flat in that region. The latter inversion is sometimes found when stimuli in this region are equated for brightness (Honig, Thomas, & Guttman, 1959).

The direction of preference within each stimulus pair agrees with the difference between the number of responses given to the members of that pair in the SS test. Accordingly, the stimulus nearer 550 mμ, received more responses in all pairs except 490, 510 mμ and 560, 570 mμ. It also appears that the degree of preference is systematically related to the difference in response level indicated for the corresponding stimuli on the SSG.

The DSG is very similar to the SSG except that the central values are somewhat higher. The "predicted DSG values" are derived from a method of predicting response totals in the DS situation which will be presented below.

Experiment 2.—The SSG and DSG are so similar to those obtained in

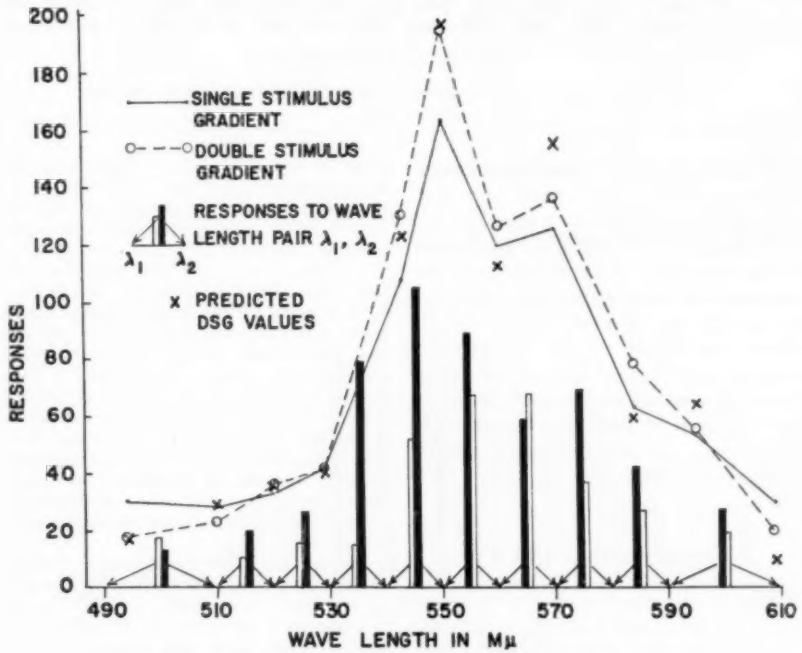


FIG. 1. Single-stimulus and double-stimulus gradients and responses to wave-length pairs, Exp. 1.

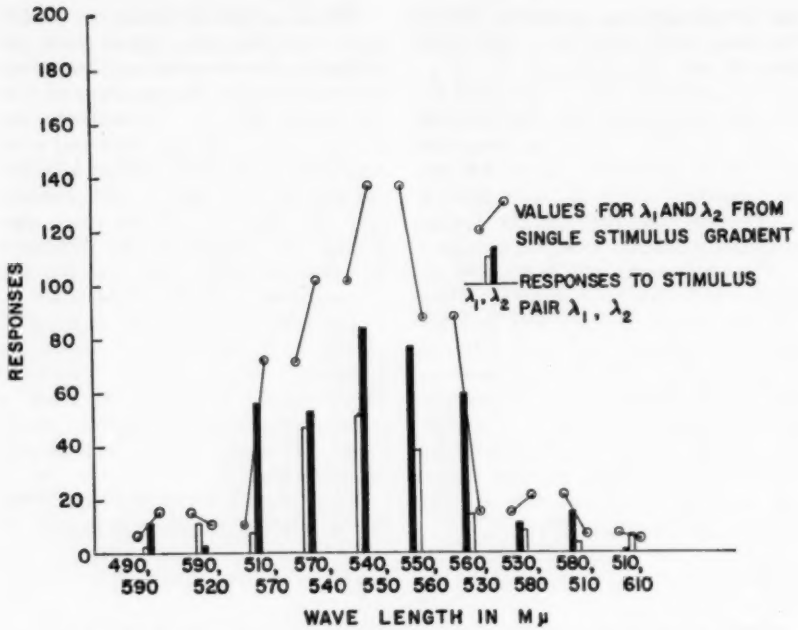


FIG. 2. Responses to wave-length pairs with corresponding single-stimulus values, Exp. 2.

Exp. 1, except for a somewhat lower response level, that these are not separately presented on a graph.

A comparison between the responses to the members of each pair and the responses to the corresponding stimuli on the SS test is presented in Fig. 2, where the latter values are plotted over the appropriate bars. The direction of preference within each stimulus pair is again in good agreement with SS values. In only two pairs is there a reversal: 530, 580 mμ and 510, 610 mμ.

EXPERIMENTS 3 AND 4

Method

Experiments 3 and 4 provide data on the effects of successive and simultaneous discrimination training on the SSG and DSG and on the distribution of responses between pairs of stimulus values. Aside from the discrimination procedure, the method in these experiments was similar to that for Exp. 1. The same apparatus and 14 of the same *S*s were used.

Procedure.—On the day following the completion of the generalization test in Exp. 1, *S* received 10 continuous reinforcements with one key illuminated in order to reinstate the conditioned operant. Discrimination training then began immediately. Each session consisted of 30 1-min. periods of stimulus presentation alternating with 10 sec. of blackout. Responding to 550 mμ (S+) was reinforced on the VI schedule used previously. Responding to 560 mμ (S−) was never reinforced, and the reinforcement programmer was interrupted when it was presented in Exp. 3.

In the *successive* discrimination training of Exp. 3, both keys were illuminated either by S+ or S− during each period. S+ and S− were each presented for 15 periods in a session. The order was randomized but excluded the presentation of the same condition for more than three consecutive periods.

Reinforcement was available at the left key for half of the S+ periods and at the right key for the other half.

The criterion for discrimination was reached when a block of 10 consecutive 1-min. periods was completed under the following conditions: (*a*) The block contained five positive and five negative periods. (*b*) At least two of the negative periods were consecutive. (*c*) *S* gave no responses to S−. (*d*) *S* gave 10 or more responses during each presentation of S+. All except 1 *S* reached this criterion within 10 sessions. A second generalization test (see below) was administered to this *S* under the considerations that on Training Sessions 7 through 10, less than 2% of its total responses were to S−, with no responses to S− on as many as four successive S− periods.

In the *simultaneous* discrimination training of Exp. 4, one key was illuminated by S+ and the other by S− during each period. The side was randomized, with the exclusion of more than three successive periods with either arrangement. The criterion for discrimination was reached when a block of five consecutive 1-min. periods was completed under the following conditions: (*a*) S− appeared on one key for two periods and on the other key for three periods. (*b*) *S* gave no responses to S−. (*c*) *S* gave at least 10 responses each period to S+. Of the 8 *S*s in Exp. 4, 6 reached criterion within five sessions or less. The other 2 *S*s stabilized at about 20% responding to S−. They apparently developed a chained response of pecking at S− and then S+ in rapid succession. This could be reinforced, since no delay of reinforcement was contingent on responding to S−. These 2 *S*s were therefore discarded.

Generalization testing.—When *S* reached the criterion for discrimination, the training session was discontinued. On the 2 following days, *S* was given a generalization test identical to the one it received after VI training in Exp. 1.

Results

A full description of the course of discrimination training in these experiments has been presented elsewhere (Honig, 1958). Suffice it to say that it differed for the two studies

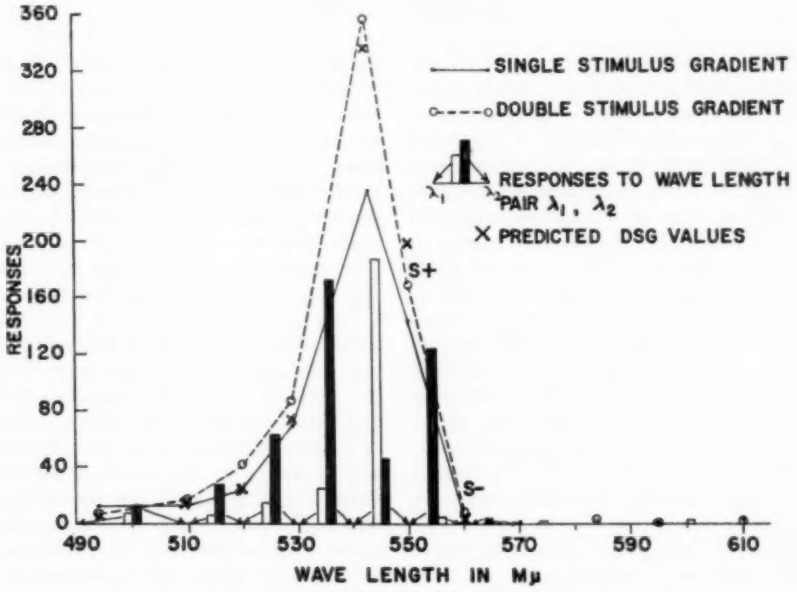


FIG. 3. Single-stimulus and double-stimulus gradients and responses to wave-length pairs, Exp. 3.

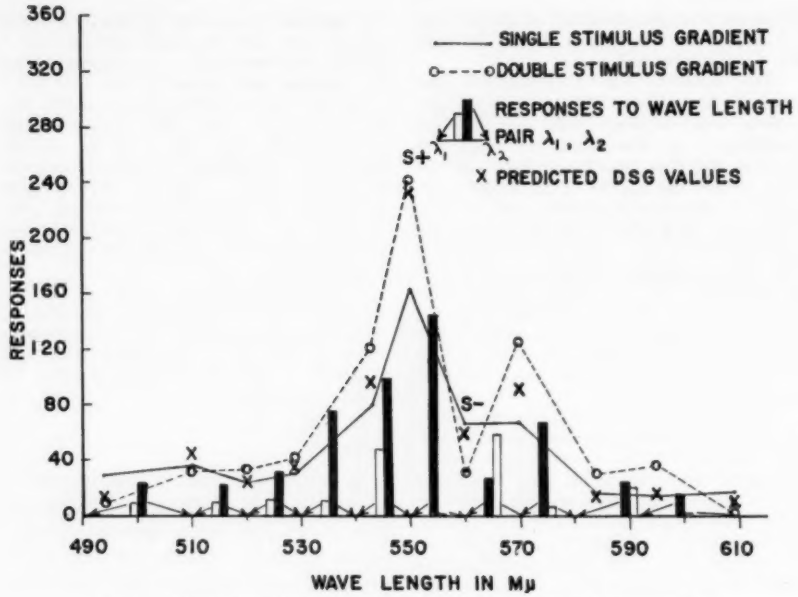


FIG. 4. Single-stimulus and double-stimulus gradients and responses to wave-length pairs, Exp. 4.

in a number of respects: (*a*) The rate of responding to S+ was much higher in the successive discrimination. (*b*) The median number of responses to extinction for S− was seven times as great in the successive discrimination. (*c*) The proportion of responses to S− was also consistently higher in that condition. The time to reach criterion was twice as great in Exp. 3, but since S+ and S− were each available only half the time in the successive procedure, the training time in terms of total minutes of presentation of each stimulus was about the same.

Single-stimulus gradients.—The differences between the SSGs following the two training procedures may be seen by comparing Fig. 3 and 4. The gradient obtained after successive discrimination entirely confirms that obtained by Hanson (1959) in that (*a*) the response level at wave lengths above 550 mμ is virtually zero; (*b*) the mode of the gradient is not at 550 mμ, but at 540 mμ; (*c*) the slope of the gradient is steeper on both sides of the mode than after simple acquisition.

The SSG obtained in Exp. 4 does not differ radically in form from the postacquisition gradient of Exp. 1. The mode remains at 550 mμ, and there is considerable responding to values between 560 and 610 mμ, with a small inversion between 560 and 570 mμ. It does appear, however, that the level of responding between 560 and 590 mμ is reduced. The number of responses obtained on the SS test is 32% less than that obtained for the same 6 *S*s in Exp. 1 for the values below 550 mμ, and 57% less for the values above 550 mμ. The largest reduction (a mean of 75 responses, or 54%) was at 560 mμ, and the reductions (in terms of absolute amount)

decreased with only one inversion between 560 and 610 mμ. All *Ss* showed more reduction at the values above 550 mμ than below.

Responding within wave-length pairs: Exp. 3.—The direction of preference for all the stimulus values conforms to the SSG. This is especially significant for the pair 540, 550 mμ, as the direction of preference is toward the postdiscrimination peak of 540 mμ. The preference of 540 mμ over S+ is an instance of *transposition of discrimination.* For the pairs comprised of the values between 490 and 540 mμ, the direction of preference is clearly toward S+. These are instances of *transposition-reversal.*

Experiment 4.—The direction of preference conforms to the SSG in all but two pairs: 510, 520 mμ and 590, 610 mμ. Both of these pairs are comprised of values near the ends of the gradient, where the slope is essentially zero. There was no transposition between 540 and 550 mμ, quite in accordance with the absence of a peak shift. Responding to the negative stimulus presented alone on the SS test was not extinguished. The simultaneous discrimination did not transfer completely to the successive situation, while the successive discrimination transferred almost perfectly to the simultaneous situation, as shown by the bars for the 550, 560 mμ pair in Fig. 3.

Double-stimulus gradients.—The DSGs are in both experiments quite similar to the SSGs though on the average somewhat higher, following the pattern set in Exp. 1 and 2. At the modes of the gradients the DSG points are considerably higher, and there are also large differences for Exp. 4 at 560 and 570 mμ. These differences can be understood on the basis of the analysis of DS rates to be presented in the next section.

Prediction of Double-Stimulus Values

Given the total responses R_1 to λ_1 and R_2 to λ_2 on a SS test, the total responses r_1 and r_2 for the same values in the pair λ_1, λ_2 on a DS test can quite adequately be predicted by the formulae:

$$r_1 = R_1\left(\frac{R_1}{R_1 + R_2}\right)$$

and

$$r_2 = R_2\left(\frac{R_2}{R_1 + R_2}\right)$$

Predictions from these functions have been carried out for each *S* for each DS value in each experiment to obtain predicted values for r_1, r_2, and $r_1 + r_2$. Mean predicted are plotted against mean obtained values for each stimulus pair from Exp. 1 and 2 in Fig. 5. The three groups of predicted values, r_1, r_2, and $r_1 + r_2$, are separated along the vertical axis for clarity. The diagonal lines represent perfect prediction ($x = y$); deviations from this are indicated by the vertical distance between each point and the line. Identical analyses carried out for Exp. 3 and 4 are not presented

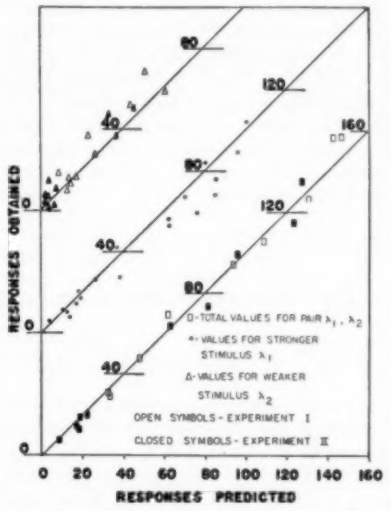


FIG. 5. Predicted and obtained double-stimulus values, Exp. 1 and 2.

here as the general outcome was very similar to that presented for Exp. 1 and 2.

From Fig. 5 it is evident that this method leads to satisfactory results. A goodness of fit test suggested by Lindquist (1953, pp. 344 ff.) was used to test the hypothesis that the obtained do not differ significantly from the predicted values. The error was divided into two components, "vertical placement," or differences between predicted and obtained means, and "departure from pattern," or the error remaining after these means have been equated. Each component provides a mean square which can be divided by the error variance of the obtained data to obtain an *F* ratio when the data have been corrected for systematic effects, in this case wave length and *S*s. This analysis was carried out for the sets of r_1 and r_2 values for the four experiments. The only errors reaching significance are for r_2 in Exp. 1, where both vertical placement and departure from pattern are significant at the .001 level. This one obtained significant departure from prediction indicates that by the present method, r_2 is systematically underestimated following simple acquisition.

Prediction of DSG values.—Consider the three stimulus values λ_1, λ_2, and λ_3, with corresponding SS response totals R_1, R_2, and R_3. A predicted DSG value for λ_2 (r_{2DSG}) is obtained when the predicted values to λ_2 are summed across the pairs λ_1, λ_2, and λ_2, λ_3. The two DS response values to λ_2 (r_{2a} and r_{2b}) are predicted as follows:

$$r_{2a} = R_2\left(\frac{R_2}{R_1 + R_2}\right) \quad r_{2b} = R_2\left(\frac{R_2}{R_2 + R_3}\right)$$

Adding these and collecting terms:

$$r_{2DSG} = R_2\left(\frac{R_2}{R_1 + R_2} + \frac{R_2}{R_2 + R_3}\right)$$

Each of the fractions can run from 0 to 1. Therefore, the coefficient for R_2 can run from 0 to 2, and thus $0 \leqslant r_{2DSG} \leqslant 2R_2$. The relative size of r_{2DSG} and R_2 (the corresponding obtained SSG value) depends on the sum of the two fractions, and this in turn depends on the shape of the gradient at λ_1, λ_2, and λ_3. In the case, for example, where λ_2 is a peak, each of the fractions must be greater than .5 since R_2 is greater than R_1 and R_3; the predicted DSG value must therefore be higher than the SSG value. In general, the two fractions will add to 1 (thus providing equal SSG and predicted DSG values) only if the points on the gradient form a geometric series.[3] If the gradient is rising in linear fashion (arithmetic series) the predicted DSG points will lie above the SSG values. The predicted DSG values are indicated as crosses in Fig. 1, 3, and 4 for Exp. 1, 3, and 4. A comparison with the obtained values supports the present analysis; particularly where differences between the DSGs and SSGs are predicted, they are also obtained. The DSG peaks, both predicted and obtained, are considerably higher than the SSG peaks, and for the central values, where the gradients tend to be linear rather than geometric, the DSG lies above the SSG.

Discussion

The foregoing analysis indicates that the direction and degree of preference between stimuli lying on the generalization gradient can be predicted from the independent response strengths of the stimuli. The occurrence of transposition in conjunction with a shift in the mode of the gradient follows from the general predictive principles outlined above. In this respect, Spence's (1937) analysis of transposition and his prediction of transposition-reversal have been supported. On the other hand, the difference in results between the successive discrimination, which produced a peak shift and transposition, and the simultaneous, which did not, are not anticipated by Spence. While equivalent criteria were demanded at the end of discrimination training under both procedures, in that *S* did not respond to S− for five successive periods, it appears from Fig. 4 that simultaneous discrimi-

[3] This can be demonstrated algebraically by a proof not given here.

nation training did not result in extinction to S− presented alone. In this discrimination, it was only necessary for S− to become a cue for switching to S+ for perfect discrimination to be manifested. In the successive discrimination, of course, *S* had no alternative response to S− available during the negative periods to obtain reinforcements, which resulted in extinction under that stimulus condition.

It appears that a shift in the peak of the PDG and the concomitant transposition depended in the present study on the development of genuine extinction to S− alone, rather than a preference for S+. But the critical factor responsible for this extinction need not have been in the manner of presentation of the discriminanda; there is nothing about the simultaneous presentation of S+ and S− per se that necessarily prevented transposition. To evaluate the factor of the manner of stimulus presentation, it would be necessary to develop a procedure where all other aspects of the simultaneous and successive problems are kept equivalent, such as the reinforcement schedules, the alternatives available for responding to S−, and so forth. Thompson (1955) and Riley, Ring, and Thomas (1960) have in effect done this by making both stimuli available on each learning trial, but preventing the animal from comparing the appropriate discriminanda in the "successive" case. They obtained more transposition when a comparison was possible. Whether the same would hold true for the pigeon working with the spectral dimension is an open question.

While the arithmetic model presented above for the prediction of DS from SS values is satisfactory for that purpose, it is not derived from assumptions about, or observations on, specific behavior patterns in the choice situations. It is possible, however, to identify the terms entering the predictive formulae with basic response dimensions in the choice situation, and thus to suggest a pattern of behavior that can be verified. Assume that in the choice situation the pigeon responds to each stimulus value at the same *rate* as when that stimulus is presented alone. This would result in a DS response total of R_1 to λ_1 if there were no interference from the presence of the other choice stimulus. Assume further that the other stimulus does interfere in that its presence reduces the total *duration* of responding to λ_1, and that the total duration of responding to both values is divided in a proportion reflecting the rates to each value. The proportion for λ_1 would be $\frac{R_1}{R_1 + R_2}$, which, when multiplied by R_1, provides the formula used above for the prediction of response totals to λ_1. The corresponding expression for λ_2 can be similarly derived. This analysis rests, of course, on a clear distinction between operant rate and the duration of the application of a given rate; while this distinction has been supported empirically by Gilbert (1958), it awaits direct confirmation in the present circumstances.

Summary

The relationship between stimulus preference and the response strength of singly presented stimuli was investigated with the use of the generalization gradient to provide stimuli of different strengths. After being trained to peck at a 550-mμ stimulus on a VI schedule, pigeons were given two concurrent generalization tests: one consisting of single stimulus values ranging from 490 to 610 mμ, and one consisting of pairs of such values. The direction of preference within each pair was found to be in direct accordance with the number of responses obtained on the single values during the generalization test. Two groups of birds then received discrimination training between 550 mμ as S+ and 560 mμ as S−, one with successive stimulus presentations and one with simultaneous. After this, the generalization and preference tests were administered a second time. The successive discrimination produced a shift in the mode of the gradient away from S− and a concurrent transposition of the discrimination between S+ and the new mode, with transposition-reversal beyond that mode. The simultaneous discrimination produced little change in the gradient and no concurrent transposition.

An arithmetic model is proposed for the prediction of responses to members of stimulus pairs from single-stimulus values. The determination of stimulus preference by the generalization gradient follows directly, and transposition is seen to be no more than a special case. The differential results obtained from successive and simultaneous discriminations are discussed with reference to the actual extinction to S− obtained from each procedure. The operant situation is analyzed in terms of some simple response dimensions to provide a basis for the arithmetic model from which double-stimulus values are predicted.

References

BAKER, R. A., & LAWRENCE, D. H. The differential effects of simultaneous and successive stimulus presentation on transposition. *J. comp. physiol. Psychol.*, 1951, **44**, 378–382.

BLOUGH, D. S. Spectral sensitivity in the pigeon. *J. Opt. Soc. Amer.*, 1957, **47**, 827–833.

EHRENFREUND, D. A study of the transposition gradient. *J. exp. Psychol.*, 1952, **43**, 81–87.

GILBERT, T. F. Fundamental dimensional properties of the operant. *Psychol. Rev.*, 1958, **65**, 272–282.

GRICE, G. R. Visual discrimination learning with simultaneous and successive presentation of stimuli. *J. comp. physiol. Psychol.*, 1949, **42**, 365–375.

GUTTMAN, N., & KALISH, H. I. Discriminability and stimulus generalization. *J. exp. Psychol.*, 1956, **51**, 79–88.

HANSON, H. M. Effects of discrimination training on stimulus generalization. *J. exp. Psychol.*, 1959, **58**, 321–334.

HONIG, W. K. Prediction of preference, transposition, and transposition-reversal from the generalization gradient. Unpublished doctoral dissertation, Duke University, 1958.

HONIG, W. K., THOMAS, D. R., & GUTTMAN, N. Differential effects of continuous extinction and discrimination training on the generalization gradient. *J. exp. Psychol.*, 1959, **58**, 145–152.

KENDLER, T. S. An experimental investigation of transposition as a function of the difference between training and test stimuli. *J. exp. Psychol.*, 1950, **40**, 552–562.

LINDQUIST, M. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.

RILEY, D. A., RING, K., & THOMAS, J. The effect of stimulus comparison on discrimination learning and transposition. *J. comp. physiol. Psychol.*, 1960, **54**, 415–421.

SPENCE, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, **44**, 430–444.

THOMPSON, R. Transposition in the white rat as a function of stimulus comparison. *J. exp. Psychol.*, 1955, **50**, 185–190.

(Received July 23, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 249–252

RESISTANCE TO EXTINCTION AS A FUNCTION OF AGE AND SCHEDULES OF REINFORCEMENT [1]

NORMAN KASS [2]

University of Minnesota

The purpose of this study is to investigate developmental changes in resistance to extinction as a function of the percentage of reinforcement provided during acquisition. Relatively little information is available concerning the effects of partial reinforcement on extinction with children, and no studies investigating age changes have been reported.

There is reason to assume that the development of cognitive processes and the growth of experience with increasing age would result in increasingly rapid recognition of the change in reinforcement schedule between an acquisition and extinction period in a simple operant task, with a resulting decrease in the resistance to extinction following partial reinforcement. Subjects at four age levels were tested, and in order to increase the generality of the findings, four partial reinforcement schedules were employed. Additional groups received 0% and 100% reinforcement.

METHOD

Subjects.—The *S*s were 216 preschool and elementary school boys and girls, approximately evenly divided by sex. The *S*s were obtained from the Institute of Child Development Nursery School, the Village Nursery School, and the Elementary School of the University of Minnesota; and from a public elementary school in Minneapolis. All schools enrolled *S*s of approximately the same intellectual and socioeconomic levels. All *S*s of the relevant age levels in attendance at the time of testing were used in each school.

Apparatus.—The apparatus was a simulated slot machine, consisting of a blue wooden cabinet with dimensions of 12 × 14 × 12 in. Centered on the front of the cabinet and 3½ in. from the top was a circular opening 1 in. in diameter. A large number of pennies was visible behind a piece of clear plastic covering the opening. The box containing the pennies was illuminated with a 7-w. lamp. Pennies were dispensed through a rectangular opening located on the lower front of the cabinet. A ½-in. wooden shaft, to which a 3 × ½ in. aluminum handle was attached, projected from the front of the cabinet 3 in. from the right and 6¼ in. from the bottom.

The rear of the cabinet was open to allow *E* to operate the apparatus. A 12 × 14 in. board was attached to each side of the machine to prohibit *S* from observing *E* load the machine.

Attached to the inside base of the cabinet was a dispensing device, which consisted of a wheel 11 in. in diameter, with 24 slots circling the outside edge. The wheel was attached to the base by a pivot with a spring-loaded brake which allowed the wheel to turn without spinning freely. This wheel was rotated one slot at a time by means of a pawl. The pawl rode in a track and was driven by a lever attached to the wooden shaft described above. When *S* pulled the handle, it rotated the wheel one slot forward and also activated a counter to record the response. On the base of the cabinet was an opening over which the wheel rotated with dimensions corresponding to those of the bottom of each slot. This opening was connected by means of a chute to the opening in the front of the machine. The pennies placed in the various slots could be dispensed when a particular slot was in line with the opening in the base. With this arrangement any variable-ratio schedule could be presented. Each machine had to be reloaded after it was played 24 times.

[1] This study was conducted while the author was a National Institute of Mental Health Postdoctoral Fellow at the Institute of Child Development, University of Minnesota, 1960–61. The writer wishes to express his appreciation of the cooperation given in this study by the principal and teachers at Kenwood Elementary School, Minneapolis, and at the University of Minnesota laboratory schools.

[2] Now at San Diego State College.

The apparatus was extremely simple to operate, and involved pulling the handle down and releasing it. When the handle was pulled down, the light which illuminated the pennies went out and relit when the handle was returned to its original position.

At the completion of the experiment *Ss* traded their pennies for prizes, such as toy wrist watches, badges, umbrellas, flutes, and kaleidoscopes.

Design.—The variables employed were percentage of reinforcement and chronological age of *S*. The six reinforcement schedules provided reinforcement on 0%, 16⅔%, 33⅓%, 60%, 80%, and 100% of the acquisition trials. The *Ss* were selected from four age levels: 4, 6, 8, and 11 yr.

Each *S* was given an acquisition and an extinction series. The acquisition series consisted of 30 trials. During this period *S* received, according to his reinforcement schedule, either 0, 5, 10, 18, 24, or 30 reinforcements. The reinforcements were presented randomly with the restriction that all *Ss* except those in the 0% condition were reinforced on the last trial of the acquisition series.

The study thus employed a 4 × 6 design with 9 *Ss* in a cell. The *Ss* at each age level were assigned to the reinforcement conditions at random.

Procedure.—The *Ss* were obtained by *E* from the classroom and taken to an experimental room located in a quiet section of the building. The *S* was told that he was going to play the "penny machine game" and was seated before the apparatus. The pennies in the illuminated window were pointed out by *E*.

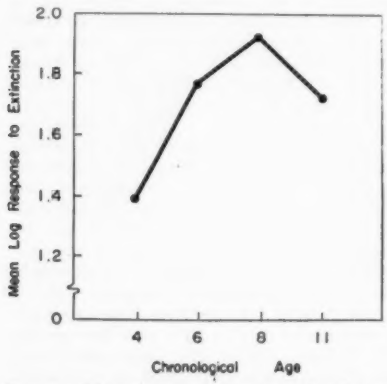


FIG. 1. Mean log responses to extinction as a function of CA of *Ss*.

The *E* said:

> You see this machine is full of pennies. Do you see all these pennies? Now in this game you try to get as many pennies out of the machine as you can, because the more pennies you get out of the machine, the better the prize you can buy with the pennies.

The *E* showed *S* the prizes he could buy. The *E* continued:

> See all the prizes we have that you can buy with the pennies you get out of the machine. Now I'll show you how to play the game. As long as the light is on, you can pull the handle and the pennies come out of here. You can play the game as long as you want. You don't have to hurry. Just tell me when you want to stop. Remember, the more pennies you get, the better the prize you can buy with your pennies.

The last statement was repeated on Trials 10 and 25 of the acquisition series. There was no indication from *E* when the acquisition series ended and the extinction series began.

Comments and questions from *S* were ignored by *E*, and if *S* persisted, *E* replied, "You can play the game as long as you like. Tell me when you want to stop playing."

Each *S* played until he indicated he wanted to stop or after 370 extinction trials, when *S* was stopped by *E*. Only 8 *Ss* failed to complete the 30 acquisition trials and were replaced. All of these *Ss* were at the 4-yr. age level and were in either the 0% or 100% group.

At the end of the session *S* was allowed to choose a prize to buy with his pennies. The *Ss* in the 0% group were told that even though they had not won any pennies, they could have a prize for playing so well. The *Ss* were asked not to tell other children about the game.

RESULTS AND DISCUSSION

The primary score used in the analysis of the results was the number of responses made during the extinction period. Because of the heterogeneity of variance among the cells, a log transformation of the scores was performed.

The results related to age differences are shown in Fig. 1, which presents the mean log of the number of responses made during the extinc-

tion period at each CA level for all percentages of reinforcement combined.

The results related to percentage of reinforcement are shown in Fig. 2, which presents the mean log responses to extinction for each percentage of reinforcement for all age levels combined. The results are in line with those obtained with college *Ss* by Lewis and Duncan (1956, 1957). There was a consistent decrease in number of responses to extinction with increasing percentage of reinforcement from $16\frac{2}{3}\%$ to 100%. The *Ss* in the 0% group made somewhat fewer responses during extinction than *Ss* in the $16\frac{2}{3}\%$ group, but made a greater number of responses than *Ss* in any other group.

In analysis of variance of the transformed scores, differences associated with CA were highly significant ($F = 9.54$, $df = 3/192$, $P < .001$), as were the differences associated with percentage of reinforcement ($F = 15.65$, $df = 5/192$, $P < .001$). The interaction term was not significant ($F = .45$).

The Duncan multiple range test (Duncan, 1955) was used to determine which of the age levels and percentages of reinforcement differed significantly. The scores of *Ss* at CA 4 differ significantly ($P < .01$) from those of *Ss* at the other CA levels; the scores of the remaining groups do not differ significantly from each other. The following differences according to percentage of reinforcement are significant at the .01 level: $16\frac{2}{3}\%$ vs. each of the larger percentages; 0% vs. 60%, 80%, 100%; 33.3% vs. 80% and 100%.

The results may be interpreted by the hypothesis that the rate of extinction is decreased as the discrimination between the acquisition and extinction series becomes more difficult. The maximum similarity between acquisition and extinction trials occurs in the 0% group, where there is no way for *Ss* to discriminate the end of the first series and the beginning of the second. The maximum dissimilarity occurs in the 100% group. The slight but nonsignificant increase from the 0% to the $16\frac{2}{3}\%$ group provides the only deviation from a consecutively decreasing number of trials to extinction with increasing percentage of reinforcement. The lack of a significant interaction between percentage of reinforcement and CA indicates that, in general, the trend of the results is similar at each age level. There was a curvilinear relationship found between CA and number of responses to extinction. These differences do not agree with those which would be expected from the hypothesis that increasing age results in an increasing ability to discriminate between changes in patterns of reinforcement. According to such an hypothesis the older *Ss* would be expected to show the least resistance to extinction. Since the only significant differences were those found between *Ss* at CA 4 and *Ss* at higher CA levels, the main problem is to determine why the youngest *Ss* should extinguish more readily. This finding might be explained in terms of the developmental changes in length of attention span. Preschool children may

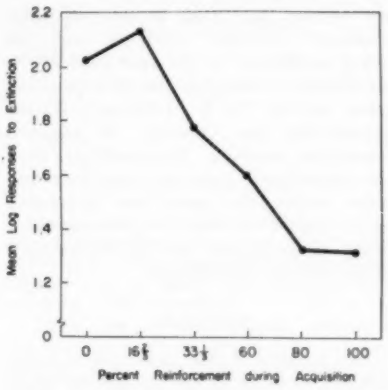


FIG. 2. Mean log responses to extinction as a function of percentage of reinforcement during the acquisition series.

not persist in a task as long as older children. Further evidence for this effect is offered by the fact that all *S*s who failed to complete the 30 acquisition trials were at CA 4. Thus even during acquisition the younger *S*s stopped responding sooner. This indicates that the extinction differences found may only reflect some other more basic difference which would be manifest over a wider range of conditions than is represented in the present experiment.

Summary

The purpose of this study was to determine the effects of six different percentages of reinforcement upon extinction of a lever-pulling response. Four age levels were employed to determine developmental changes in resistance to extinction. The apparatus employed was a simulated slot machine designed for use with young children.

A total of 216 children at CA 4, 6, 8, and 11 received reinforcement on either 0, $\frac{1}{6}$, $\frac{1}{3}$, $\frac{2}{3}$, $\frac{5}{6}$, or all of the trials of a 30-trial acquisition period. The extinction trials were continued until *S* wished to stop or until 370 extinction responses had been made. The response measure was the total number of responses to extinction.

A decrease in number of responses during extinction was found with increasing percentages of reinforcement. The least number of trials to extinction was shown by *S*s at CA 4 and the greatest by *S*s at CA 8. There was no significant interaction between CA and percentage of reinforcement.

References

Duncan, D. B. Multiple range test and multiple *F* tests. *Biometrics*, 1955, **11**, 1–42.

Lewis, D. J., & Duncan, C. P. Effect of different percentages of money reward on extinction of a lever-pulling response. *J. exp. Psychol.*, 1956, **52**, 23–27.

Lewis, D. J., & Duncan, C. P. Expectation and resistance to extinction of a lever-pulling response as functions of percentage of reinforcement and amount of reward. *J. exp. Psychol.*, 1957, **54**, 115–120.

(Received July 24, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 253–257

SUBJECTIVE SCALE OF FORCE FOR A LARGE MUSCLE GROUP[1]

HANNES EISLER[2]

Harvard University

Although subjective force is one of the classical continua studied in psychophysics, most studies of force have been limited to lifted weights. Recently, however, Stevens and Mack (1959) scaled subjective force, as exerted in the squeezing of a handle. Their experiments confirmed the psychophysical power law; the exponent obtained was 1.7. Borg and Dahlström (1960) investigated muscular work carried out on a bicycle ergometer. Though the variable they studied was power rather than force, their experiments can also be included in investigations of muscular effort. They found an exponent of 1.6.

The investigation reported here consists of a series of experiments carried out to scale the subjective force exerted by a comparatively large muscle group. The muscle group chosen contains those muscles that are used in pushing or pressing a pedal with the foot in a horizontal, forward direction. In addition to being intrinsically interesting, these experiments make it possible to compare the subjective scale of force exerted by the large leg muscles with the scale obtained for the smaller muscles of the forearm used in squeezing a hand dynamometer. To summarize the results, a power function was obtained for foot pressure also, and its exponent did not differ appreciably from the exponent for handgrip.

[1] This research was supported by a grant from the National Science Foundation (Psycho-Acoustic Laboratory Report No. PNR-260).

[2] Now at Psychology Department, University of Stockholm, Sweden.

General Method

Scaling.—Five methods were used: magnitude estimation, magnitude production, matching of force of handgrip to force of foot pressure and the reverse, and cross-modality matching of both handgrip and foot pressure to white noise. The magnitude estimation was repeated once, and the matching of foot pressure to handgrip twice with slight variations, so that, all in all, eight experiments enter into the present study. Except in the first magnitude estimation experiment in which *O* gave four judgments for every stimulus, two values for every stimulus were obtained from each *O* in all experiments.

Subjects.—Twelve *O*s were used in each experiment. The groups largely overlapped and were identical for the first magnitude estimations and the magnitude productions, as well as for all the experiments in which foot pressure and handgrip were matched. Altogether, 16 men and 5 women took part in the investigation, mostly graduate students of psychology and staff members.

Apparatus.—In all the experiments, *O* sat in a rigid chair and with his right foot pressed a pedal in a forward, horizontal direction. The distance through which the pedal moved was very small.

In the experiments with magnitude estimation and the matching of handgrip to foot pressure, the pedal was connected through a lever system to the platform of a large beam scale. The measuring beam of the scale was loaded with a given *stimulus* weight, and *O*, by pressing the pedal, brought the scale into equilibrium. Equilibrium was indicated to him by a pointer on a display.

In the experiment in which force of foot pressure was matched to force of handgrip, a handle was connected with the platform of the scale. Forces exerted as *responses* (foot pressure in magnitude production, foot pressure when matched to handgrip, etc.) were measured with tensile gauges connected with the pedal, the handle, or both, depending on the particular experiment.

When force of foot pressure and force of handgrip were matched to white noise, *O*s listened to the band of noise (75 to 2400 cps) through a pair of earphones.

Sources of error.—Three sources of error were evident in the series of experiments reported, though not every error was to be found in every experiment:

1. The pointer in the display could not be kept completely stationary. There seemed to be some correlation between the speed of its movement and the force exerted. This movement may or may not have influenced *O*'s judgment. Since this source of error did not exist in all the experiments and since the agreement among all the experiments is quite good, it is probably of only minor importance.

2. People differ in strength. The problem caused by this state of affairs could be attacked by (*a*) limiting the range investigated to the maximum force of the weakest *O*, but restriction of range is not an appealing solution; (*b*) picking a particular sample of strong *O*s, but restriction of sample is not an attractive solution either; (*c*) or dropping *O*s as the forces required were increased. This solution was chosen. It is a legitimate solution if the exponent of the power function is independent of the strength of *O*. In that case the only effect would be a decrease in the reliability for strong stimuli, since the points obtained are based on fewer *O*s.

3. While he pushed the pedal, *O* had to hold his leg up by his own force. For small stimuli the force required was apparently greater than the force of pushing. For large stimuli the friction between the sole of the shoe and the pedal appeared to be sufficient to eliminate this extra force. The *O*s seemed to be trying to neglect the holding force and only to take the pushing force into account.

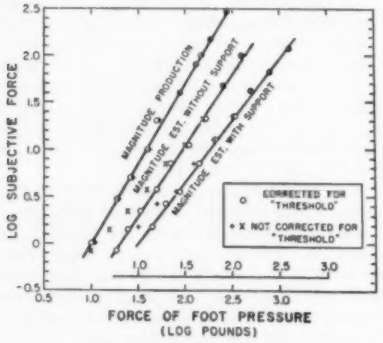


FIG. 1. Subjective force of foot pressure in logarithmic units as a function of physical force in log pounds. (The upper scale on the abscissa refers to the magnitude estimation experiment in which *O*'s leg was supported.)

MAGNITUDE ESTIMATION AND MAGNITUDE PRODUCTION

Method

Whether the function obtained by magnitude estimation was independent of the strength of *O*s was tested crudely in two ways. (*a*) The *O*s were divided into a group of 7 who were able to exert the greatest force used, 400 lb., and a second group of 5 who were not. Medians of the magnitude estimates of the six lowest stimuli common to the two groups were computed for each group separately. A plot of the medians of the two groups against each other was essentially linear. (*b*) Medians were computed, for each group separately, of the magnitude estimates for the highest stimulus force each *O* could exert, the second highest stimulus presented, and so on. This procedure is meaningful when the stimuli are spaced logarithmically, on the assumption that the power law holds for this case. Again, when the values for one group are plotted against those for the other, the result is essentially linear. Thus the function relating subjective force to physical force appears to be largely independent of the strength of the particular *O*.

The fact that a plot of the logarithms of subjective force (geometric means) against the logarithms of force in pounds was somewhat curved indicates that the "threshold" cannot be neglected. Consequently, the values were corrected for "effective threshold" or "subjective zero," by a method developed by Ekman (1961).

Results

The data of the magnitude estimation experiment indicated that subjective force is a power function by Ekman's criterion. In contrast to most other scaling experiments, however, this one gave a *negative* threshold. The threshold correction amounted to +8.9 lb. The logarithms of the magnitude estimates are plotted in Fig. 1 against the logarithms of the corrected (circles) and uncorrected (crosses) physical values. A least squares fit yielded a slope, that is to say, an exponent of 1.51.

The magnitude production data were treated in the same way as the magnitude estimation data. For mag-

nitude production, the threshold was positive and the correction small, −1.8 lb. The data are also shown in Fig. 1. The exponent was 1.70.

It is common to find this discrepancy between the exponents for magnitude estimation and magnitude production; S. S. Stevens has called it the "regression effect" (see e.g., Reynolds & Stevens, 1960). The *O* decreases the range of the variable over which he has control—numbers in magnitude estimation and forces in magnitude production.

The consensus of these two experiments is that, after application of threshold corrections, subjective force of foot pressure grows as a power function of physical force with an exponent of approximately 1.6. The result agrees with the findings of Borg and Dahlström (1960), derived from fractionation experiments. They too obtained the exponent 1.6 (and a negative threshold).

In order to find out whether the negative threshold found in the magnitude estimation experiment was related to the third source of error mentioned above—the force required to hold up the leg—the magnitude estimation experiment was repeated with the variation that *O*'s heel was supported by a leather belt. The results are given in Fig. 1. The negative threshold was confirmed, although the threshold correction was smaller in this experiment, +3.8 lb. The exponent, however, decreased to 1.31. Since *O*s were divided about whether the supporting belt was an improvement, and some of them found it more uncomfortable than the first experiment, the results of this experiment will not be taken into account in the conclusion of the whole series of experiments.

Matching Experiments

Method

The procedure used in the experiment in which handgrip was matched to foot pressure was similar to that of the experiment with magnitude estimation, except that, instead of responding with numbers, *O* now equated squeezes on a handle with the force exerted by pressing the pedal.

Matching foot pressure to handgrip is the inverse of the procedure described above. Three variations were carried out. Two of them differed from each other only in the stimuli chosen and the leverage applied between handle and scale. The leverage was changed in such a way as to increase the precision of the forces presented by *E*. In the third variation the *O*s were asked to press the pedal, not simultaneously with the hand squeezes, but after them.

In order that the results for simultaneous and successive matching could be compared, the medians of the two experiments were plotted against each other. Since a straight line was obtained, indicating that successive matching made no difference, the results from the experiment with successive matching were not treated further.

In the final experiment, *O*s were asked to match in random succession both force of foot pressure and force of handgrip to a band of white noise, whose intensity varied between 30 and 100 db. re 0.0002 dyne/cm^2. The point of this experiment was to match force of foot pressure with force of handgrip indirectly. A noise of a certain intensity was presented and *O* was informed whether he was to press the pedal or to squeeze the handle. This was the only experiment in which the weakest *O*s did not drop out as the forces were increased.

Results

All the matching experiments gave curvilinear functions when force of handgrip was plotted against force of foot pressure in log-log coordinates (geometric means). (For the experiment with white noise, force of handgrip matched to a certain noise intensity was plotted against the force of foot pressure matched to the same noise intensity.) Plotted linearly, however, straight lines (that did not pass through the origin) were obtained. This outcome indicated (*a*) a nonzero threshold for at least one of the continua (with no way to tease out the separate threshold corrections), and (*b*) a power relation between subjective force of foot pressure and subjective force of handgrip, with an exponent of 1. Thus, if either of the

two continua grows as a power function of physical magnitude, the other grows with the same exponent.

The straight lines were computed according to a variation of the method of least squares. The customary method weights the high values more than the low values. Because of the comparatively great ranges covered, the application of that method would have introduced a heavy bias. Therefore the sum of squares of the *relative* deviations,

$$\Sigma[(y-y_1)/y]^2$$

where y equals experimental values and y_1 computed ones, has been minimized rather than the sum of squares of the *absolute* deviations,

$$\Sigma(y-y_1)^2.$$

The lines obtained were normalized so that they passed through the origin with a slope of 1, and they are summarized in Fig. 2. Figure 3 shows the outcome of the individual matching experiments in log-log plots after the threshold correction has been arbitrarily carried out on the foot pressure continuum. The lines drawn in Fig. 3 have a slope of 1.

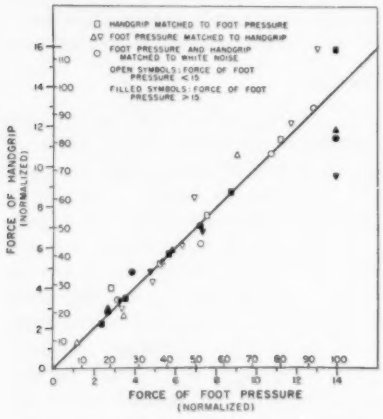

□ HANDGRIP MATCHED TO FOOT PRESSURE
ΔΩ FOOT PRESSURE MATCHED TO HANDGRIP
○ FOOT PRESSURE AND HANDGRIP MATCHED TO WHITE NOISE
OPEN SYMBOLS: FORCE OF FOOT PRESSURE <10
FILLED SYMBOLS: FORCE OF FOOT PRESSURE >10
FORCE OF HANDGRIP (NORMALIZED)
FORCE OF FOOT PRESSURE (NORMALIZED)

FIG. 2. Force of handgrip as a function of force of foot pressure from four matching experiments. (The values are normalized such that the four different straight lines obtained by a variation of the method of least squares are made to coincide as one line, which passes through the origin and has a slope of 1. Because the range covered is great, it is divided into two sections which are superimposed.)

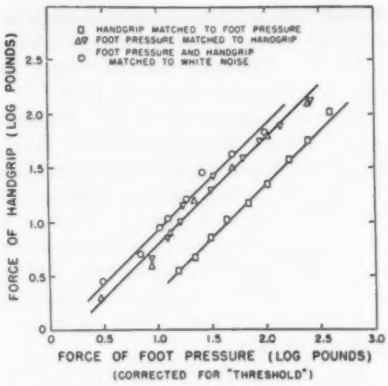

□ HANDGRIP MATCHED TO FOOT PRESSURE
ΔΩ FOOT PRESSURE MATCHED TO HANDGRIP
○ FOOT PRESSURE AND HANDGRIP MATCHED TO WHITE NOISE
FORCE OF HANDGRIP (LOG POUNDS)
FORCE OF FOOT PRESSURE (LOG POUNDS)
(CORRECTED FOR "THRESHOLD")

FIG. 3. Force of handgrip (in log pounds) as a function of force of foot pressure (in log pounds) from four matching experiments. (The values of foot pressure are corrected for "threshold.")

It is of interest to notice that the parameter k in Stevens' power law

$$\psi = k\phi^n$$

where ψ refers to a subjective magnitude and ϕ to a physical magnitude, seems to depend on the particular experimental procedure. In matching, two sets of subjective magnitudes are equated (Stevens, 1959):

$$\psi_1 = k_1\phi_1^{n_1}$$
$$\psi_2 = k_2\phi_2^{n_2}.$$

In the present case, since $n_h = n_f$, this yields

$$\phi_h = (k_f/k_h)\phi_f$$

where the subscripts h and f refer to handgrip and foot pressure. It is the factor k_f/k_h whose logarithm constitutes the intercepts in Fig. 3. Note that the data of both experiments in which foot pressure was matched to handgrip scatter

around the same straight line, whereas the other two matching experiments yield different intercepts. Some kind of rule of least effort seems to hold, more conspicuously for foot pressure: the forces over which O has control are small relative to those presented by E, thus yielding a large coefficient k; when both forces are matched to white noise, the ratio k_f/k_h becomes almost 1.

Summary

The subjective force of pushing a pedal with the leg has been scaled as an instance of the subjective force exerted by a large muscle group. The following methods were employed: magnitude estimation, magnitude production, matching the force of handgrip to the force of foot pressure and vise versa, and matching both foot pressure and handgrip to the intensity of white noise. The experiments involving numbers yielded a power function with an exponent of 1.6 relating subjective force to physical force. All the matching experiments showed that the exponent for force of foot pressure and force of handgrip is the same. The exponent for handgrip has previously been determined as 1.7. Thus the subjective force of foot pressure, as measured in this study, approximates a power function of physical force not appreciably different from the one found earlier for force of handgrip. The exponent for foot pressure approximates 1.65.

References

Borg, G., & Dahlström, U. The perception of muscular work. Publication No. 5, 1960, Umeå Research Library, Sweden.

Ekman, G. A simple method for fitting psychophysical power functions. *J. Psychol.*, 1961, **51,** 343–350.

Reynolds, G. S., & Stevens, S. S. The binaural summation of loudness. *J. Acoust. Soc. Amer.*, 1960, **32,** 1337–1344.

Stevens, J. C., & Mack, J. D. Scales of apparent force. *J. exp. Psychol.*, 1959, **58,** 405–413.

Stevens, S. S. Cross-modality validation of subjective scales for loudness, vibration, and electric shock. *J. exp. Psychol.*, 1959, **57,** 201–209.

(Received July 27, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 258–266

RE-EXAMINATION OF THE SERIAL POSITION EFFECT[1]

MURRAY GLANZER

Department of Psychiatry, University of Maryland

AND STANLEY C. PETERS[2]

Walter Reed Army Institute of Research

Existing generalizations about the serial position effect in rote learning have recently been challenged as based on artifacts. McCrary and Hunter (1953) presented evidence that the different curves produced by experimental variables, such as presentation rate, distribution, and list difficulty, are a single curve multiplied by different numbers of errors. Their contention was borne out by the findings of Braun and Heymann (1958). A re-examination of the serial position effect seems necessary. Since any statement about serial position effects must refer to the beginning and end of a series of items, the first step in the re-examination is to analyze the meaning of the terms *beginning* and *end*. The analysis consists of three steps: to determine what characteristics *E* points to as defining the beginning and the end of a list, to analyze these characteristics, and to determine their effect on *S*'s performance.

In a rote learning experiment, *E* presents to *S* a continuous and repetitive cycle of events: the series of syllables, a gap; the series of syllables, a gap. The terms *beginning* and *end* are used to refer to the characteristics associated with this gap. They are the following: (*a*) Primacy-recency: The first item *S* sees, and the last item he sees before the cycle repeats itself, appear on either side of the gap, because *E*s usually present the list starting from the gap. (*b*) Spacing: The appearance of the gap coincides with a period in which *S* is not required to anticipate syllables. (*c*) Association break: Every item in the series is both a stimulus and a response item, except the two on either side of the gap. The *S* is not required to form an association across the gap. The first item in a list is usually a pure stimulus item. The last item is a pure response item, in that it does not function as stimulus for the first item.

The purpose of the following experiments was to separate the factors that make up the complex called *the beginning* and *the end* of the list, and to determine the part these factors play in the serial position effect.

EXPERIMENT I: PRIMACY-RECENCY

Mitchell's (1934) evidence that the usual bowed curve appears in lists that have neither spacing nor association breaks would indicate a considerable role for primacy-recency effects. In Exp. I, the role of primacy-recency was evaluated by varying *S*s' starting positions: starting some *S*s at the normal position in the list, and starting other *S*s at what would ordinarily be called the middle of the list. The normal position in the list is that part which has the gap followed by the asterisk, or other starting cue; this may be called the *structural beginning*. The *temporal*

[1] This work was carried out under Contract DA-49-007-MD-1004 between the Office of The Surgeon General and the University of Maryland.

[2] Now at the University of Illinois.

beginning, a term which may be used to refer to the first syllable exposed to *S*, may or may not coincide with the structural beginning. The two experimental conditions here were one in which the structural and temporal beginning coincided, and one in which they did not coincide. If the temporal starting point (the primacy-recency factor) plays an important role in determining the serial position effect, then groups starting at the structural middle should show a flattened or indented serial position curve.

Method

Subjects.—The *S*s were 32 Army medical service enlisted men. They were average or above average in their scores on Army intelligence tests (General Technical score of the Army Classification Battery).

Materials.—Two six-syllable lists (Glaze 20% association value) were used: List I: *, ZID, NUK, WEF, QAM, TUH, BEJ; List II: *, QAM, TUH, BEJ, ZID, NUK, WEF. The two lists consist of the same series of syllables, with the structural beginning placed at different positions.

Procedure.—The *S*s were first given two three-syllable practice lists (Glaze 100% association value). One practice list was started from the structural beginning, i.e., the asterisk. The other was started from the structural middle. Half the *S*s were started from the structural beginning in their first practice list; the other half were started from the structural middle in their first practice list. After learning both practice lists to a criterion of one perfect trial, *S* learned one of the six-syllable lists.

The presentation rate was 3 sec. per syllable, with a 6-sec. interval (two blank spaces on the drum) between successive cycles of the list. All groups received the same instructions: to try to anticipate each syllable before it appeared. No reference was made to either the beginning or the end of the list. The list was learned to a criterion of three consecutive errorless trials. Half the *S*s learned List I; the other half, List II. Half the *S*s started at the structural beginning with the asterisk exposed. The other half started at the structural middle of the list, with either WEF or BEJ exposed. Trials were counted as starting at the structural middle of the list for the *S*s who started at that point.

Results

Since the work of McCrary and Hunter (1953) and Braun and Heymann (1958) showed that the total number of errors has a multiplicative effect on the shape of the serial position curve, here, and in the subsequent experiments, error scores were translated into logarithms. This translation converts the multiplicative effect to an additive factor and permits direct comparison of the shapes of the curves. This procedure was adopted instead of dividing through by the total number of errors, as in the work mentioned above.

The groups that started in the structural middle had a slightly higher serial position curve (see Fig. 1) than those starting at the structural beginning. The overall differences between the two experimental groups

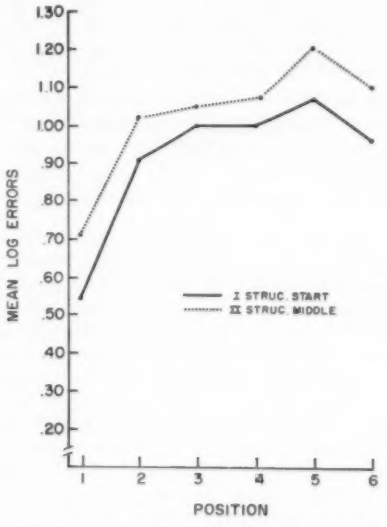


FIG. 1. Serial position curves with primacy-recency varied: Exp. I. (Group II started in the middle of the list.)

TABLE 1

ANALYSIS OF VARIANCE OF LOG ERROR SCORES: EXP. I

Source	*df*	*MS*	*F*
Between *Ss*			
Condition (C)	1	.5645	1.69
Lists (L)	1	.3934	
C × L	1	.1068	
Error (between)	28	.3331	
Within *Ss*			
Position (P)	5	.9834	42.76***
P × C	5	.0171	
P × L	5	.0193	
P × C × L	5	.0384	1.67
Error (within)	140	.0230	

*** $P < .001$.

were not, however, statistically significant (see Table 1). For the group starting at the asterisk, the overall mean of the converted error scores was 5.5 ($SD = 1.4$); the mean number of trials to criterion was 19.7 ($SD = 9.2$). For the group starting at the structural middle, the mean of the converted errors was 6.1 ($SD = 1.4$); the mean number of trials to criterion was 22.9 ($SD = 9.4$). The experimental operation had no discernible effect on the shape of the serial position curves, and of course the Position × Experimental Condition interaction was not significant. List structure was the sole effective factor, significant at the .001 level.

Greenhouse and Geisser (1959) have pointed out that the significance levels used in analyses of variance in repeated measurements designs are incorrect, if the assumption of equal covariances is not met. In that case, the usual test, presented in Table 1, overestimates the significance level. A lower bound test that underestimates the significance level can be obtained by appropriate reduction of the degrees of freedom. The effect of list structure remains significant at the .001 level ($df = 1/28$).

The lists used in this experiment were shorter than the 10- to 12-syllable lists usually used in serial position studies. However, the curves (Fig. 1) show the same skewed bow shape as the curves obtained with longer lists (Hovland, 1938a). They are also similar to the family of curves obtained by Robinson and Brown (1926) for lists that range from 5 to 17 syllables. There is no basis, therefore, for assuming that different factors are at work in 6-syllable lists than those in longer lists.

The type of scoring used above (based on the total number of errors before the criterion was met) may conceal an early effect of primacy-recency that is swamped by the repeated effect of list structure. To check on this possibility, the number of the first trial on which a correct anticipation occurred was scored for each position in the list. With the starting position at the asterisk, the means for the six positions were 2.38, 5.00, 8.06, 8.50, 9.75, and 6.62. With the starting position in the middle of the list, the means were 3.81, 6.38, 6.31, 7.81, 11.31, and 8.50. There is an indication of flattening of the curve for the lists started at the middle. The Position × Condition interaction, however, was not significant ($df = 5/140$, $P > .10$).

It is clear that the structure of the list, rather than a primacy-recency effect, is the major determinant of the serial position effect; the same serial position effect appears with or without primacy-recency. The results do not contradict Mitchell's (1934) results, since in her lists the structural characteristics of spacing and association break were absent. As will be pointed out below, when the structural characteristics are removed, primacy-recency effects can be seen. For the usual type of list presentation, however, the important factor seems to be the structure of the list. The next step was to examine the two factors associated with the structural beginning and end of the list: spacing and the association break.

EXPERIMENT II: SPACING

If spacing is a factor determining the appearance of serial position effect, then changing the amount of spacing

should affect the shape of the serial position curve. More specifically, if the confounding multiplicative effect of total number of errors is eliminated, then increasing the spacing should give more pronounced serial position effect. McCrary and Hunter (1953) and Braun and Heymann (1958) have presented evidence that contradicts this hypothesis, but they compared only 6-sec. and 126-sec. intertrial intervals. Intervals between 0 and 16 sec. were included in the experiment below. Perhaps more important than the range of spacing values is the fact that in the Hovland (1938b, 1940) studies re-analyzed by McCrary and Hunter, and also in the Braun and Heymann study, the long and short spacing intervals were not completely comparable. In those studies, an interpolated task was given during the 126-sec. interval, but not during the 6-sec. interval.

Method

Subjects.—The *S*s were 90 Army medical service enlisted men. The group was above average on Army intelligence tests (General Technical score on the Army Classification Battery).

Materials.—Two 10-syllable (Glaze 20% association value) lists were used: List I: *, RUW, GIY, POH, WEF, QAM, ZIX, NUK, BEJ, XOC, KAZ; List II: *, ZIX, NUK, BEJ, XOC, KAZ, RUW, GIY, POH, WEF, QAM.

Procedure.—The *S*s were assigned in equal numbers to each of the two list conditions and the five spacing conditions. They first learned two 3-syllable practice lists (Glaze 100% value) to a criterion of one perfect trial, under the same spacing condition as their main list. They then learned one of the 10-syllable lists to a criterion of two errorless trials. The criterion was set as high as possible within the time available, so that the serial position curves would represent final, stabilized performance.

The lists were presented at a 2-sec. rate, with zero, one, two, four, or eight blank spaces (0, 2, 4, 8, or 16 sec.) between the end of the list and the reappearance of the beginning. There were 9 *S*s who were unable to reach the criterion within the 75 min. available: 2 each in the zero- and one-space condition, 1 each in the two- and eight-space condition, and 3 in the four-space condition. These *S*s were replaced by others.

Results

The mean log error curves show an orderly change from a relatively flat curve for the zero-spacing condition, to a sharply peaked curve for the maximum spacing condition (see Fig. 2). The statistical evaluation (see Table 2) of the effect with the usual tests, however, does not find the Spacing × Position interaction significant ($.10 > P > .05$). This analysis does not, of course, reflect the orderly nature of the changes in the curves. For example, there is a progressive increase in the slope of the curve between Positions 1 and 6 as the spacing increases. The rank order correlations for the difference between Positions 1 and 6 with the amount of spacing is 1.00 for both the List I and List II groups. In both groups, there is a corresponding, but less marked, decrease in slope between

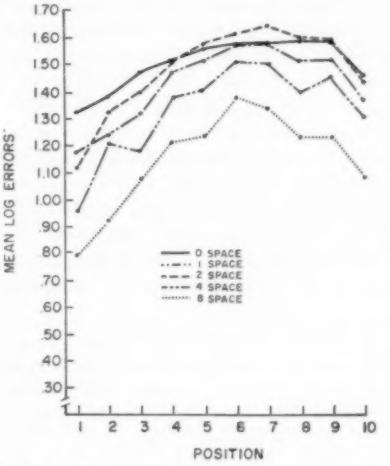


FIG. 2. Serial position curves with spacing varied: Exp. II.

TABLE 2

ANALYSIS OF VARIANCE OF LOG ERROR SCORES: EXP. II

Source	*df*	*MS*	*F*
Between *Ss*			
Spacing (S)	4	3.5962	7.00***
Lists (L)	1	0.0059	
S × L	4	0.3396	
Error (between)	80	0.5138	
Within *Ss*			
Position (P)	9	2.0380	90.18***
P × S	36	0.0315	1.39
P × L	9	0.0437	1.93*
P × S × L	36	0.0168	
Error (within)	720	0.0226	

* $P < .05$.
*** $P < .001$.

Positions 6 and 10 as spacing increases. The rank order correlation for the difference between Positions 6 and 10 with the amount of spacing, is .70 in both groups.

The analysis demonstrates the expected overall differences between the spacing conditions, the means declining ($P < .001$) with increased spacing, except for a reversal between the one- and two-space conditions. The overall means of the converted errors for the zero- to eight-space conditions, in order, were as follows: 15.0 ($SD = 2.1$), 14.3 ($SD = 2.1$), 14.8 ($SD = 1.9$), 13.4 ($SD = 2.7$), 11.6 ($SD = 2.2$). The means of the number of trials to reach criterion were, in order, as follows: 58.4 ($SD = 21.8$), 54.4 ($SD = 19.6$), 60.3 ($SD = 21.4$), 47.7 ($SD = 22.1$), 35.0 ($SD = 18.0$).[3] The Position × Lists interaction reflects differences in the difficulty of the syllables. The absence of a significant third-order interaction indicates that the effect of the spacing on the serial position curve does not differ for the two lists. The position effect is significant ($P < .001$) and remains at that level under the Greenhouse-Geisser lower bound test.

The results lent some support to the hypothesis that the serial position curve changes in an orderly fashion as a function of spacing. The statistical support was, however, not satisfactory.

[3] The contribution of intrusion errors to the total error score may be of interest. The number of complete intrusions was tallied and expressed as a percentage of total errors for each *S*. The mean percentage of intrusion errors for the zero- to eight-space conditions, in order, was as follows: 5.3 ($SD = 6.1$), 4.7 ($SD = 6.5$), 3.0 ($SD = 2.5$), 4.2 ($SD = 2.7$), 7.7 ($SD = 6.7$).

EXPERIMENT III: SPACING AND ASSOCIATION BREAK

In this experiment, the aim was to retest the effect of spacing on the serial position effect and also to evaluate the effect of the association break. The expectation was that spacing would increase, and that associative chaining across the gap would decrease, the serial position effect.

Method

Subjects.—The *Ss* were 120 college students. They were paid for their participation.

Materials.—In order to insert associative chaining between successive presentations of the list, the lists were constructed in quasi-paired-associates format. Ten chained lists were constructed by rotating the syllables systematically through the 10 available positions. Two of the chained lists are shown below.

Chained 1: * ZIX NUK * NUK BEJ * BEJ XOC * XOC KAZ * KAZ RUW * RUW GIY * GIY POH * POH WEF * WEF QAM * QAM ZIX (Space 0, 2, or 8) * ZIX NUK * (etc.)

Chained 2: * NUK BEJ * BEJ XOC * XOC KAZ * KAZ RUW * RUW GIY * GIY POH * POH WEF * WEF QAM * QAM ZIX * ZIX NUK (Space 0, 2, or 8) * NUK BEJ * (etc.)

To break the chaining, the initial syllable in each of the 10 chained lists was replaced with a substitute syllable, i.e., LEB for ZIX in Chained 1; CIW for NUK in Chained 2; and TIV for BEJ, HUQ for XOC, VEP for KAZ, FOJ for RUW, LEQ for GIY, JAT for POH, YUS for

WEF, and FUP for QAM, in the other 8 chained lists.

All syllables, including the substitute syllables, were from the Glaze 20% association value list.

Procedure.—The presentation rate was 1 sec. per syllable. Three spacing conditions were used: 0, 2, and 8 spaces, or seconds, between lists. The 10 lists, 3 spacing conditions, and 2 chaining conditions formed a $10 \times 3 \times 2$ factorial design. Two *Ss* were assigned to each of the 60 experimental conditions.

The drum displayed an asterisk, followed by a stimulus syllable and then the response syllable. A stimulus syllable, therefore, appeared every 3 sec. After learning two 3-syllable practice lists in the same quasipaired-associates format to a criterion of one errorless trial, each *S* learned a 10-syllable list to a criterion of two successive errorless trials.

Results

Spacing has the expected effect of decreasing total errors during learning. The effect is significant at the .01 level (see Table 3). The overall means of the converted errors for the zero-, two-, and eight-space conditions, in order, were as follows: 13.6 ($SD = 2.1$), 13.0 ($SD = 2.0$), 12.0 ($SD = 1.7$). The means of the number of trials to reach criterion were, in order, as follows: 44.7 ($SD = 18.2$), 39.6 ($SD = 15.0$), 31.9 ($SD = 11.1$). Chaining does not have any significant overall effect. The serial position effect was significant ($P < .001$), with the Greenhouse-Geisser lower bound test leaving the significance level unchanged.

TABLE 3

ANALYSIS OF VARIANCE OF LOG ERROR SCORES: EXP. III

Source	*df*	*MS*	*F*
Between *Ss*			
Spacing (S)	2	2.5434	6.09**
Chaining (C)	1	0.0097	
Lists (L)	9	0.3966	
S × C	2	0.2680	
S × L	18	0.3114	
C × L	9	0.7316	1.75
S × C × L	18	0.2184	
Error (between)	60	0.4175	
Within *Ss*			
Position (P)	9	1.0312	56.35***
P × S	18	0.0609	3.33***
P × C	9	0.0250	1.37
P × L	81	0.0584	3.19***
P × S × C	18	0.0261	1.43
P × S × L	162	0.0172	
P × C × L	81	0.0178	
P × S × C × L	162	0.0184	1.01
Error (within)	540	0.0183	

** $P < .01$.
*** $P < .001$.

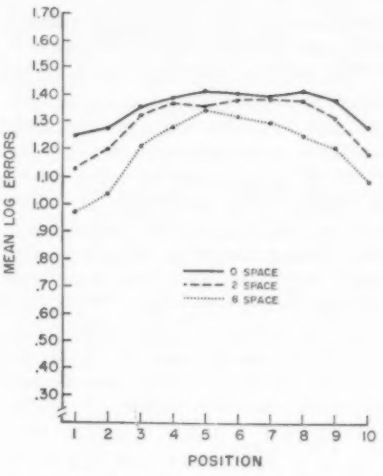


FIG. 3. Serial position curves with spacing varied: Exp. III.

The effect of spacing on the serial position curve is shown in Fig. 3. As spacing increases, the curve moves toward a markedly bowed shape of classic form, with the beginning lower than the end. The effect, as evaluated by the Position × Spacing interaction, is significant at the .001 level. Use of the Greenhouse-Geisser lower bound test with 2 and 60 *df* leaves the effect significant, at worst, at the .05 level. The effect of chaining on the serial position curve was not significant. The significant interaction of Lists and Serial Position ($P < .001$) reflects differences in the difficulty of

individual syllables, which were in different positions in the 10 lists.

The plots of the mean unconverted error scores display the same progression toward a more peaked serial position curve, with increased spacing. Plots of percentages of total errors, of course, give essentially the same picture, with respect to the shapes of the curves, as that obtained with mean log of errors. The data were also converted into ranks. Thus, if *S* had the following raw error scores for the 10 positions: 6, 8, 15, 21, 18, 22, 19, 24, 13, 17, his scores were converted to the following ranks: 1.0, 2.0, 4.0, 8.0, 6.0, 9.0, 7.0, 10.0, 3.0, 5.0. Examination of the mean ranks for each position showed the same progression found with the mean log scores. Using the ranks, coefficients of concordance were computed for each group. For both the chained and unchained groups, the coefficients increased as spacing increased. In the chained group, they were .18, .30, .48; in the unchained group, they were .11, .42, .48. All the coefficients are significant ($P < .05$). The regular progression of the coefficients indicates that the increase in the bowing of the serial position curve as a function of spacing is independent of the particular conversion used.

Figure 3 shows a slight serial position effect in the zero-space condition. Separate analysis of the zero-space condition finds the effect significant ($F = 7.39$, $df = 1/60$, $P < .01$). The chained zero-space condition by itself also demonstrates a significant serial position effect ($F = 4.94$, $df = 1/60$, $P < .05$). The factor underlying this effect is the primacy-recency variable. Evidently, as in Mitchell's (1934) study, when spacing is absent, the primacy-recency factor affects the shape of the curve. The effect is, however, too slight to account for the usual serial position effect.

Discussion

The preceding experiments indicate that of the three factors listed initially as possible determinants of the serial position effect (chaining, spacing, and primacy-recency), the most important factor is spacing, the appearance of a gap in the list. Experiment I eliminated primacy-recency as a major factor in the usual serial learning situation, since a full, unaffected serial position effect was shown even when opposed by the primacy-recency factor. Experiment II gave some evidence that the serial position effect was a function of spacing. Experiment III completed the case in support of the effect of spacing and eliminated chaining as a major determinant of the serial position effect. Experiment III also indicated a slight effect for primacy-recency when it is not opposed by spacing.

McCrary and Hunter (1943), in their re-analysis of Hovland's (1938b, 1940) data, and Braun and Heymann (1958) display curves that remain unaffected by spacing, once the multiplicative effect of total number of errors is eliminated. Evidence has been presented here that spacing does have an effect. The apparent contradiction probably stems from differences in procedure. In the previous studies, the intertrial intervals were 6 and 126 sec.; in the present experiments, they ranged between 0 and 16 sec., with large changes appearing between the 0- and 2-sec. conditions. In the previous studies, an interpolated task was used, but only during the long interval. This could have counteracted the spacing effect by slowing the learning of the beginning and end syllables of the long interval lists.

Inhibition (Hull, Hovland, Ross, Hall, Perkins, & Fitch, 1940) and interference (Atkinson, 1957) constructs have been popular in the explanation of the serial position curve. From one point of view, the results above fit in with inhibition or interference explanations of the serial

position effect. Spacing could function as a barrier to protect the items near the beginning and the end of the list. With the demonstration of a systematic effect of spacing on the serial position curve, however, another type of explanation becomes possible: a facilitation explanation. With this type of explanation, interference or inhibitory effects between list items are considered homogeneous within the list. The serial position curve is viewed as a result of the facilitative effect of spacing on the learning of the first and last items of the list. These, in turn, facilitate the learning of their neighbors. The skewing of the curve can be explained by an effect demonstrated by Ribback and Underwood (1950). They showed that once an association is learned between a pair of syllables, it is easier to attach a third association to the second member of the pair than to the first. In other words, once the association X-Y is learned between syllables X and Y, it is easier to learn the triplet of syllables X-Y-Z than it is to learn A-X-Y. The Ribback-Underwood mechanism by itself is not sufficient to generate the serial position curve, since the mechanism does not determine the syllables from which it will spread. This determination could be made by spacing, which would facilitate the syllable pairs bordering the gap. These anchor syllables would in turn generate the serial position effect forward and backward from the gap.

Summary

Serial position effects are usually defined on the basis of the beginning of a list. In rote learning, the term "beginning" may be analyzed into three factors associated with a repetitively-appearing gap that separates the "end" from the "beginning" of the list: (*a*) Primacy-recency. The first item the *S* sees, and the last item he sees before the cycle repeats itself, appear on either side of the gap, because *E*s usually start the list from the gap. (*b*) Spacing. The appearance of the gap coincides with a period in which the *S* is not required to anticipate syllables. (*c*) Chaining versus association break. Every item in the series functions as both a stimulus and a response item (chaining), except the two on either side of the gap. One of these is solely a stimulus item; the other is solely a response item.

Three experiments were carried out to determine the role of these three factors in generating the serial position effect. In Exp. I, primacy-recency was opposed to both spacing and association break by varying *S*'s starting position in conventional lists. No effect of the variation was found, indicating that primacy-recency was not a major factor. In Exp. II, spacing was varied, using intertrial intervals of 0, 2, 4, 8, and 16 sec. With the multiplicative effect of total number of errors held constant, some evidence was obtained indicating that the serial position curve became more peaked as spacing increased. In Exp. III, both spacing (intertrial intervals of 0, 2, and 8 sec.) and chaining or association break were varied. To vary chaining, the lists were given in quasipaired-associates form. Some lists required that an association be formed between the end and the beginning of the list; other lists did not. Experiment III showed clearly the effect of spacing in determining the serial position effect. There was no evidence for the effect of association break or chaining. Experiment III also indicated a slight effect for primacy-recency when it is not opposed by spacing. On the basis of these findings, it is concluded that the major factor determining the serial position effect is the amount of space between the end and the beginning of the list, with an increase in spacing producing a more marked serial position effect.

References

Atkinson, R. C. A stochastic model for rote serial learning. *Psychometrika*, 1957, **22**, 87–96.

Braun, H. W., & Heymann, S. P. Meaningfulness of material, distribution of practice, and serial-position curves. *J. exp. Psychol.*, 1958, **56**, 146–150.

Greenhouse, S. W., & Geisser, S. On methods in the analysis of profile data. *Psychometrika*, 1959, **24**, 95–112.

Hovland, C. I. Experimental studies in rote-learning theory: II. Reminiscence with varying speeds of syllable presentation. *J. exp. Psychol.*, 1938, **22**, 338–353. (a)

Hovland, C. I. Experimental studies in rote-learning theory: III. Distribution of practice with varying speeds of syllable

presentation. *J. exp. Psychol.*, 1938, **23**, 172–190. (b)

HOVLAND, C. I. Experimental studies in rote-learning theory: VII. Distribution of practice with varying lengths of list. *J. exp. Psychol.*, 1940, **27**, 271–284.

HULL, C. L., HOVLAND, C. I., ROSS, R. T., HALL, M., PERKINS, D. T., & FITCH, F. B. *Mathematico-deductive theory of rote learning*. New Haven: Yale Univer. Press, 1940.

MCCRARY, J. W., & HUNTER, W. S. Serial position curves in verbal learning. *Science*, 1953, **117**, 131–134.

MITCHELL, M. B. The effect of serial position in the continuous memorization of numbers. *Amer. J. Psychol.*, 1934, **46**, 493–494.

RIBBACK, A., & UNDERWOOD, B. J. An empirical explanation of the skewness of the bowed serial position curve. *J. exp. Psychol.*, 1950, **40**, 329–335.

ROBINSON, E. S., & BROWN, M. A. Effect of serial position upon memorization. *Amer. J. Psychol.*, 1926, **37**, 538–552.

(Received July 27, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 267–271

EFFECTS OF INSTRUCTIONS IN PROBABILITY LEARNING [1]

J. McCRACKEN, C. OSTERHOUT, AND JAMES F. VOSS

College of Wooster

The present paper[2] reports the results of two experiments designed to investigate probability learning as a function of instructions. The major instructional aspects studied were type of problem presented and information regarding sequential characteristics of the task. Studies of the role of instructions per se (Anderson & Grant, 1957) have not been extensive. However, instructions used in studies of two-choice, noncontingent probability learning that allowed for the possibility of E_1-E_2 patterns and permitted *S* to respond on a trial-by-trial basis have yielded A_1 probabilities approximately equal to E_1 probabilities (e.g., Engler, 1958; Grant, Hake, & Hornseth, 1951; Neimark, 1956). On the other hand, studies which provided additional information regarding the nonpatterned E_1-E_2 trials yielded A_1 frequencies which exceeded E_1 probabilities (Morse & Runquist, 1960; Rubenstein, 1959). Experiment 1 was designed, therefore, to test the hypothesis that event prediction instructions and event prediction instructions with additional information regarding the sequential nature of the task yield A_1 probabilities equal to and in excess of the E_1 probability, respectively. In addition, Exp. 1 tested the hypothesis that instructions which present no problem, but tell *S* to engage in a conversation, yield nondifferential E_1-E_2 effects. A transfer condition was employed in order to determine the permanency of instructional effects.

Experiment 2 was designed to test the hypothesis that more specific instructions to avoid a trial-by-trial basis of responding and instructions to consider the task as essentially a discrimination problem yield A_1 maximization. Discrimination instructions were considered to be the extreme case of the lack of a trial-by-trial basis of responding since *S* was told to consider the E_1-E_2 trials either in blocks of trials or as one two-event discrimination.

EXPERIMENT 1

Method

The E_1-E_2 probability of .70–.30, randomized within 20-trial blocks, was used for both experiments. Each of two 33% association nonsense syllables (Hilgard, 1951), PIB and FAJ, occurred as the E_1 event for one-half of the randomly assigned *S*s. The *S*s were students at the College of Wooster.

Original training (*OT*).—Instructions for Group 1 essentially asked *S* to engage in a conversation:

> You have before you a list of two nonsense syllables. Using only the two words on this list, you and I are going to hold a conversation. To begin this conversation, you will say one of the words to me. I, in turn, will say one of the words to you. You will look over your list and say the word which you think is a reasonable statement at that time. We will continue in this manner until the end of the experiment. Remember, use only the two words before you for all of your contributions to the conversation.

[1] This research was sponsored by the National Institute of Mental Health (M-3531). The authors wish to thank N. H. Anderson for his helpful comments.

[2] The notation used in this paper is as follows: A_1 and A_2 denote *S*'s predictions of events E_1 and E_2, respectively, where E_1 is the more frequent (.70) event and E_2 the less frequent (.30) event.

Instructions for Group 2 asked *S* to predict a sequence of events:

> You have before you a card with two nonsense syllables listed on it. Using only these two words, you and I will hold a special type of conversation. You will begin by choosing one of these words and saying it aloud to me and I will then respond with one of the words. From this point on, you will try to predict the word I will say. If you are correct in your prediction, I will say that same word after you. We will continue in this manner (your prediction, my response) until the end of the experiment.
>
> Your task is merely to try very hard to predict correctly each separate word I shall say.
>
> Remember, use only the two words before you for all of your predictions and try hard to get each of your predictions right.

Instructions for Group 3 presented *S* sequential information and asked *S* to regard the situation as one problem. The first two paragraphs of the instructions for Group 2 were used, followed by:

> There is no pattern to discover. Your task is merely to get as many predictions correct as possible. Consider the entire experiment as one general problem and solve it so as to have the most correct you possibly can. It may be necessary to have some wrong predictions in order to best solve this general problem. Remember: treat the entire experiment as one general problem and attempt to solve it in order to get as many correct predictions as possible —even if you must have some wrong predictions to do so.

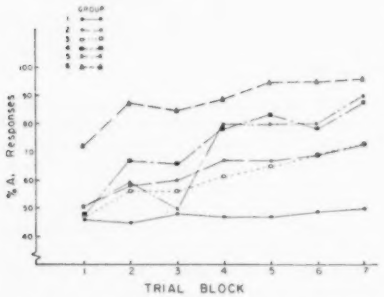


FIG. 1. Percentage of A_1 responses as a function of Trial Blocks 1–7 for Groups 1–6 of Exp. 1 and 2.

TABLE 1

PERCENTAGE OF A_1 RESPONSES FOR THE SEVEN 20-TRIAL BLOCKS OF THE NINE TRANSFER CONDITIONS

Block Trial	Group								
	11	21	31	12	22	32	13	23	33
1	44	52	50	52	60	56	48	73	69
2	51	59	58	67	73	70	63	82	76
3	56	56	61	65	76	71	73	83	74
4	46	61	56	66	76	68	78	85	75
5	50	53	59	73	75	68	80	86	78
6	49	57	63	75	70	76	84	84	83
7	48	55	59	70	76	72	76	83	85

Note.—The first digit of the group number designates the OT condition, the second designates the T condition.

Thirty *S*s served in each group for 140 OT trials. The *E* responded 2 sec. after *S* gave a response.

Transfer (*T*).—Each OT group was randomly divided into three equally sized subgroups. Each subgroup received one of the three sets of OT instructions after the 140 OT trials. The T (second instruction) groups were run for another 140 trials, with the same respective E_1-E_2 words, probability condition, and counterbalancing of words, but different 20-trial block randomized E_1-E_2 sequences.

Results[3]

Original training.—Figure 1 presents the Groups 1, 2, and 3 A_1 percentages as a function of trial block for OT trials. Group 1 responded at approximately chance level, whereas Groups 2 and 3 approached matching. An analysis of variance performed on the A_1 frequency data revealed a significant instruction group effect ($F = 32.50$, $df = 2/84$). A Duncan range test indicated 1–32.[4] A Trial

[3] A preliminary experiment was performed involving Group 1 instructions and different E_1-E_2 sequences. The results were roughly equivalent to the present Group 1 results. The effect of the verbal counterbalancing and the interactions including word used are not significant for any analysis and are not further discussed.

[4] Significant differences are indicated by dashes between the group numbers, with an increase in A_1 frequency from left to right.

Block 7 analysis of variance revealed a significant instruction group source of variation ($F = 6.89$, $df = 2/84$) and the same range results.

Transfer.—Table 1 presents the percentage A_1 responses of the nine transfer groups for Trial Blocks 8–14. The data of Fig. 1 and Table 1 indicate a relatively rapid shift in performance level with the introduction of the T instructions. Analysis of variance performed on the A_1 frequency data of Trial Blocks 8–14 revealed a significant T instruction group source of variation ($F = 25.32$, $df = 2/72$). Subsequent range test results indicated 1–2–3. The OT instruction and T Instruction × OT Instruction sources of variation are not significant.

Analyses of variance were performed on the data of Trial Blocks 8 and 14. The OT instruction group source of variation in Trial Block 8 is significant ($F = 6.38$, $df = 3/72$), with subsequent range test results of 1–23. These results agreed with the results of Trial Block 7. The OT instruction group source of variation for Trial Block 14 is not significant. The Trial Block 8 and 14 findings likely reflect the reduced influence of OT instructions.

The T instruction group source of variation is significant in the Trial Block 8 ($F = 6.49$, $df = 2/72$) and Trial Block 14 ($F = 21.02$, $df = 2/72$) analyses. Subsequent range tests revealed: 12–3 and 1–23 on Trial Block 8 and 1–23 on Trial Block 14.

Experiment 2

The lack of difference between Groups 2 and 3, except for the overall T analysis, implied that the sequential information and suggested approach presented to Group 3 did not differentiate the problem from the simple event prediction for Group 2. Experiment 2 was therefore designed to test the hypothesis that instructions to avoid a trial-by-trial basis of responding and instructions which force *S* to consider the task as essentially one two-choice discrimination problem yield maximization.

Group 4 was presented instructions to consider the task as one problem and to *avoid a trial-by-trial basis of responding*. Group 5 was instructed to predict one event when told to "Choose" (after 20 trials) and to maintain that response until again told to "Choose" (after 20 trials). Group 6 was instructed to consider the situation as a discrimination task and select the response more likely to occur once such a discrimination was made.

Method

The first paragraph of instructions for Groups 4, 5, and 6 was the same as that of Groups 2 and 3. The Group 4 instructions continued:

> Please note that you should ignore a trial-by-trial approach as much as possible. Neglect of this approach is helpful because there is no pattern or system to the sequence of words I shall say. Thus, you should approach this task as one problem and use any strategy you are able to devise in order to get as many words correct as possible. You will not be able to get all of the words correct, but you should try to get a maximum number of words correct. Remember, you are to approach the task as one problem and employ a strategy which will enable you to make a maximum number of correct predictions.

Group 5 instructions were:

> Please note that you should ignore a trial-by-trial approach as much as possible. Neglect of this approach is helpful because there is no pattern or system to the words I shall say. In addition, you will not be able to get all of the words correct. However, in order to facilitate your getting a maximum number of correct predictions, the following procedure will be used. When we begin, you will try to predict the word I shall say. Then, at a given point in the sequence, I shall say "Choose."

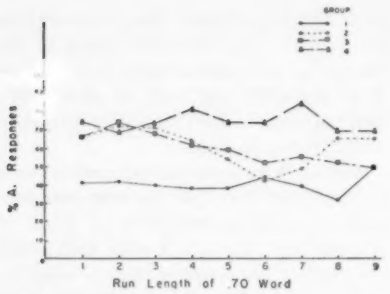


FIG. 2. Percentage of A_1 responses as a function of E_1 run length for Groups 1, 2, and 3 of Exp. 1 and Group 4 of Exp. 2.

> After I say "Choose," you are to again predict, but you must predict with the same word until I again say "Choose." This procedure will continue until the end of the experiment. Remember, when I say "Choose," you must only predict with one of the words until I again say "Choose." Naturally, you should say the word that you think will occur more frequently until I again say "Choose" so that you will make a maximum number of correct predictions.

Group 6 instructions were:

> Please note that you should ignore a trial-by-trial approach as much as possible. Neglect of this approach is helpful because there is no pattern or system to the words I shall say. In addition, I want to point out that you will not be able to get all of the words correct. However, in order to facilitate your getting a maximum number of correct predictions, you are to use the following procedure.
>
> Consider the experiment as essentially one of discrimination. Therefore, if you are able to discriminate or tell which word occurs more frequently, you are to respond with that word on all of the trials. Since there is no pattern to the words, you will get as many words correct as possible by responding in this manner. Remember, as soon as you are able to, you should tell which word occurs more frequently and respond with that word *on each prediction*.

The E_1-E_2 sequences were those of Trials 1–140 of Exp. 1. The N was 10 for each group. Other conditions were comparable to those of Exp. 1.

Results

The A_1 percentages for Groups 4, 5, and 6 during Trial Blocks 1–7 are presented in Fig. 1. A test for significance between Groups 4, 5, and 6 was not performed because of the limitations placed upon the A_1 responses by the instructions.

The results of Group 4 indicate that instructions to respond to the situation as one problem and to avoid a trial-by-trial approach yielded an A_1 percentage approaching 90%. The Group 5 A_1 percentage approached approximately the same level. Group 6 tended to choose the more frequent word during the first trial block, although after an E_2 run, an A_2 response occasionally occurred.

The differences among Groups 1, 2, 3, and 4 were clarified by an analysis of negative recency effects. Figure 2 presents the percentage of A_1 responses as a function of E_1 repetitions. Groups 1 and 4 did not show negative recency effects, whereas Groups 2 and 3 yielded such tendencies. These results imply that Groups 2 and 3 responded more to sequential characteristics of the stimuli than Groups 1 and 4. The frequency distributions of A_1 responses of the Trial Block 7 data of Groups 1, 2, 3, and 4 were essentially unimodal.

DISCUSSION

In general, the findings indicated that conversational instructions yielded chance results; event prediction instructions, with or without additional sequential information and problem orientation, yielded approximate matching; event prediction instructions which included the avoidance of a trial-by-trial orientation yielded an A_1 frequency greater than the E_1 frequency; and instructions which forced S to choose one alternative for blocks of 20 trials or to discriminate which event occurred more frequently yielded asymptotic A_1 levels of approximately 90% to 95%. The results of Group 5 are in agreement with the results of Galanter and Smith (1958).

The results are relevant to the area of

decision making in a probabilistic situation (Edwards, 1961) in that the A_1 response level is a function of the problem presented. Instructions which include the possibility of E_1-E_2 patterns and do not specifically state that a trial-by-trial approach should be avoided apparently yield results based upon sequential E_1-E_2 characteristics. On the other hand, the results of Groups 5 and 6 suggest that maximimizing occurs when the task is presented as one discrimination-type problem and the responses indicated are A_1 and A_2 for large blocks of trials rather than trial-by-trial sequences of A_1 and A_2 responses.

The conversation instruction results suggest that in a probability situation, differential E_1-E_2 frequency is not a sufficient condition to yield A_1-A_2 differences.

Summary

Instructions in two .70—.30 noncontingent probability experiments were varied for six groups. Groups 1 to 6 were told to consider the task as (*a*) conversation, (*b*) event prediction, (*c*) event prediction—with additional task information, (*d*) event prediction—avoiding a trial-by-trial basis of responding, (*e*) a problem involving responding for blocks of 20 trials, and (*f*) a problem involving the discrimination of the two events.

Group 1 responded at an approximate 50% A_1 level, Groups 2 and 3 approximately matched, and Groups 4, 5, and 6 exceeded 85% A_1 responding. Analysis of negative recency effects for Groups 1–4 suggested that Groups 1 and 4 were not responding to the sequential nature of the task, whereas Groups 2 and 3 were.

References

Anderson, N. H., & Grant, D. A. A test of statistical learning theory model for two-choice behavior with double stimulus events. *J. exp. Psychol.*, 1957, **54**, 305–317.

Edwards, W. Behavior decision theory. *Annu. Rev. Psychol.*, 1961, **12**, 473–498.

Engler, J. Marginal and conditional stimulus and response probabilities. *J. exp. Psychol.*, 1958, **55**, 303–317.

Galanter, E. H., & Smith, W. A. S. Some experiments on a simple thought problem. *Amer. J. Psychol.*, 1958, **71**, 359–366.

Grant, D. A., Hake, H. W., & Hornseth, J. P. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *J. exp. Psychol.*, 1951, **42**, 1–5.

Hilgard, E. R. Methods and procedures in the study of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 517–567.

Morse, E. B., & Runquist, W. N. Probability-matching with an unscheduled random sequence. *Amer. J. Psychol.*, 1960, **73**, 603–607

Neimark, E. D. Effects of type of nonreinforcement and number of alternative responses in two verbal conditioning situations. *J. exp. Psychol.*, 1956, **52**, 209–220.

Rubenstein, I. Some factors in probability matching. *J. exp. Psychol.*, 1959, **57**, 413–416.

(Received August 3, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 272–279

COGNITIVE FACTORS IN HEART RATE CONDITIONING [1]

BISHWA B. CHATTERJEE [2] AND CHARLES W. ERIKSEN

University of Illinois

Studies on autonomic conditioning in human *S*s typically have given little attention to the role of *S*'s cognitive or verbalizable expectancies upon the development of the conditioned response. Where attention has been directed to these variables the emphasis has been on demonstrating that humans can develop conditioned autonomic responses without verbalizable awareness. Experiments by Diven (1937) and Haggard (1943) have been interpreted as showing that human *S*s will condition GSRs without verbalizable awareness of the contingent relationship between the CS and the UCS and Razran (1946) has interpreted his data on salivary conditioning along the same lines.

In these studies, however, only a crude assessment was made of *S*'s verbalizable awareness. There was no attempt to make a fine grained analysis of specific expectancies and their relation to autonomic response. Even the conclusion of conditioning without awareness is extremely equivocal. Lacey and Smith (1954) have shown that neither the Diven nor Haggard studies contain the necessary logical or statistical comparisons of data to permit a conclusion of unconscious conditioning. Further, Chatterjee and Eriksen (1960) have shown that Lacey and Smith's own data on heart rate conditioning without awareness may well have been due to an artifact arising from the method of computing conditioning scores.

There are but few studies that have attempted to manipulate experimentally *S*'s cognitive expectancies in autonomic conditioning or to correlate closely his verbalizable expectancies with autonomic responses. Notterman, Schoenfeld, and Bersch (1952) informed one group that there would be no further shocks and found extinction of conditioned heart rate was much quicker than in an uninformed group. Branca (1957) undertook intensive questioning of his *S*s and related their expectations of shock as well as their verbalizations as to whether the shock was painful or not to the frequency and occurrence of conditioned GSRs. He concluded (Branca, 1957) "expectation of shock as a painful or fearful experience was necessary and sufficient to produce responses to the experimental and generalization stimuli in this experiment and such expectancy was the result of awareness of the existing relationships between the experimental stimuli and experience with the unconditioned stimulus" (p. 549).

In the present experiment we studied the acquisition and extinction of a conditioned heart rate response under different conditions of experimentally manipulated cognitive expectancies. Since previous methods of determining autonomic conditioning have been shown to result in spurious evidence of precise conditioning (Chatterjee & Eriksen, 1960; Eriksen, 1958) an improved and more rigorous criterion of conditioning was employed. In addition the generalization of conditioned heart rate responses along a semantic and a color dimension was determined.

[1] This research was supported by Mental Health Grant M-1206.

[2] Now at the University of Michigan.

METHOD

Design.—Stimuli were common words presented to *S* by means of an optical projector upon a ground glass screen at exposure duration of .5 sec. The *S* began chain associating to the word upon exposure. After 8.5 sec. a small light located above the screen flashed for 1 sec. This was the signal for *S* to stop associating. Following a 15-sec. rest interval, the next stimulus word was presented. The stimulus list consisted of 12 nouns and adjectives all with a high frequency count in the Thorndike-Lorge (1944) tables.

In the conditioning session the stimuli were arranged in seven blocks of trials, each block containing all 12 words randomized within the block. In the extinction phase, which followed the conditioning phase without interruption, a second list of 12 words was used. Only one word, BOAT, the CS, was common in both lists. The extinction session consisted of three blocks of trials with each block again containing randomization of the 12 words. Each word was exposed with a deep colored background of either green, yellow, or red. There were four words in each color in both stimulus lists. The extinction list contained the word BOAT, five words semantically related to BOAT, and six words semantically dissimilar. The stimulus words with the associated colors for the conditioning and extinction phases are shown in Table 1.

TABLE 1

COLOR SCHEME OF THE STIMULUS WORDS

Conditioning Stimulus List		Extinction Stimulus List	
Word	Color	Word	Color
BOAT	green	BOAT	green
CHILD	green	SEA	green
RICH	green	FACE	green
PAPER	green	INDUSTRY	green
BIRD	yellow	SHIP	yellow
HOUSE	yellow	SAIL	yellow
ALONE	yellow	DRESS	yellow
BLUE	yellow	PLOT	yellow
SOFT	red	VESSEL	red
TREE	red	WAVE	red
SKY	red	NATURE	red
PEACE	red	MONEY	red

To establish cardiac conditioning a painful shock was always presented at the termination of chain association to the word BOAT. The shock, lasting for 1 sec., coincided in time and duration with the light signal for stopping associations. There were no shocks in the extinction phase.

The electrotachographic response of *S* was continuously recorded during the experimental period. Upon completion of the extinction trials *S* was interrogated with a series of questions to estimate his awareness of shock frequency following each of the stimulus words.

Procedure.—The *S* was seated in a comfortable reclining chair facing a large wood screen that contained the ground glass projection screen. All of the apparatus was concealed from the sight of *S*. The *S* was instructed as follows:

> We're going to do an experiment on chained association. A common word will be flashed for one-half second on this screen and you are to read that word and then go on speaking the words that come to your mind. Do not try to make any meaningful sentences or any sense out of your free association. Rather, give your mind a free rein. It is best to be relaxed and not hold back anything. Continue speaking one word after another until you notice this little light bulb flash for one second. This is the signal for you to stop. Then wait until the next word is flashed on the screen when you again begin your free associations as before. We will use only a limited number of words over and over again.

The *S* was then given training with a few practice words and any questions he had were answered. Heart rate was recorded through Standard Lead 1 and shock was administered through electrodes placed on the right calf with 1½-in. interelectrode distance. The *S* was told that electric shock was an integral part of the experiment and that we would begin by determining his threshold for pain. A number of shocks were then administered at increasing voltage levels until *S* reported the shock as quite painful. This was the level used for that *S*.

Three basic groups of *S*s were formed by adding to the above instructions in the following ways. Group I *S*s were told that a shock would follow one particular word but that no other words would be followed by shock. They were further told when the extinction trials began by informing them there would be no further shocks. Group II *S*s were told that following the presentation of a particular word in the list there would always be a shock and each of the remaining words in the list

would be followed by one shock sometime during the presentation. They were further told that during the latter part of the experiment all the shocks would cease. Group III was told that a certain number of shocks would be administered at certain points of time during the experiment but *E* could not tell them beforehand when the shocks would come. These differences in instructions for the three groups were designed to lead to different cognitive expectancies concerning conditioning arrangements. In keeping with these differences in instructions Group I *Ss* were given only 7 shocks while *Ss* in Groups II and III received 18 (7 to the word BOAT and 1 each to the remaining 11 words present during the conditioning trials).

Before beginning the actual conditioning phase an adaptation period was employed in which 12 training words were used.

Upon termination of the extinction phase the electrodes were removed and all *Ss* were put through an interrogation as to their expectancies and beliefs concerning the nature of the experiment. In addition to questions designed to elicit *S*'s hypotheses as to the purpose of the experiment and the basis on which shock was administered, the list of stimulus words were administered to *S* one by one and he was asked to indicate the number of times he thought shock had followed each. This provided an estimate of *S*'s expectancy of shock for the different words.

Apparatus.—The stimulus words were typed in small letters and mounted in 35-mm. slides. Colored cellophane sheets were wrapped around the slides to provide variation in color. A Bell and Howell Robomatic slide projector was used in projecting the stimuli on the ground glass screen. Stimulus presentation, shock administration, and duration were automatically programed and timed.

A Grass Model 5P4 polygraph was used for continuous recording of *Ss*' electrotachographic responses. In addition to the pen recording the tachogram, three event marking pens were also triggered from appropriate leads from the timing equipment. They were used to indicate the onset and termination of the CS, of the stop signal, and the UCS.

Shock was faradic stimulation from a Harvard inductorium supplied with a 3-v. dc source in the primary.

Subjects.—Seventy-six undergraduate students (19 females and 57 males) in introductory psychology and education courses were used. Assignment to groups was alternated on the basis of the order of their appearance for the experimental session except that the number of females in each group was made roughly proportionate.[3]

Results

The response measure used in assessing conditioning, extinction, and generalization was a cardiac response difference score. The latencies between the four heat beats immediately prior to the presentation of the stimulus were determined for each of the 120 stimulus presentations during the conditioning and extinction trials for each *S*. Also, the latencies between the four beats immediately following presentation of the CS were determined. The cardiac response difference was obtained by subtracting the shortest post-CS latency from the shortest pre-CS latency on each presentation.[4]

The 9 *Ss* in Group I were all able to verbalize during the inquiry period that shock had followed only the critical word BOAT. The 11 *Ss* constituting Group II had received a total of 18 shocks: 7 to the word BOAT and once each to the remaining 11 words during the conditioning phase. Again all these *Ss* reported during the inquiry that they had discriminated the word BOAT from the remaining words by the third or fourth presentation of BOAT. On the other hand, none of the other 56 *Ss* comprising Group III were able to verbalize clearly and unequivocally that shock had followed BOAT nearly all the time during the conditioning trials and occurred only

[3] After 11 *Ss* had been run in each group, the data were examined. Due to recording failures 2 *Ss* had to be discarded from Group I. The existence of conditioning in Groups I and II was at this point quite evident, but the negative results in Group III led us to add an additional 45 *Ss* to this group in order to insure a more sensitive test of this condition.

[4] The cardiac UCR in the present experiment was almost without exception a decrease in latency between beats.

once to each of the remaining words. However, when *Ss* were asked to state their expectancies for shock to each of the stimulus words, it was possible to subdivide Group III into two groups: Group III-A, the partially aware group consisted of 29 *Ss* who met the criterion of reporting that they had received at least three or more shocks to the word BOAT and not more than two shocks to any of the remaining words. The remaining 27 *Ss* formed Group III-U. These *Ss* reported frequency of shock to the word BOAT less than or equal to two times or reported more than two shocks to BOAT but at the same time reported an equal or greater number of shocks to one or more of the remaining words. This division of Group III was employed to permit a more refined analysis of the relationship between cognitive expectancy and conditioning and generalization.

Conditioning and extinction.—Since conditioning is a monotonic increasing function of the number of reinforced trials, a rigorous test of conditioning is to plot the number or percent of *Ss* in each experimental group who gave the maximum positive cardiac response difference to BOAT in each block of trials (chance level is 8.3%).

The results of this measure of conditioning and extinction are plotted in Fig. 1. It is apparent that the conditioning varies from group to group. It is greatest in Group I, somewhat less so in Group II, and in Groups III-A and III-U the curves indicate that no conditioning occurred.

The binomial test was used to determine whether the obtained proportions of *Ss* in each group giving the maximum response to BOAT on the last acquisition trial block differed significantly from the chance level. For Group I the probability is .003; for Group II, .06; for Group III-A, .11; and .40 for Group III-U.

The above method for assessing conditioning is quite rigorous and is designed to reveal the specificity of the stimulus-response relation. It is informative, however, to examine conditioning and extinction in the different groups in terms of an alternative criterion. First, the sizes of the response differences in each block for each word were ranked separately from small to large for each *S* and then the rank value for BOAT was averaged through *Ss* by group and trial block. Curves obtained in this way are shown in Fig. 2.

While this measure is more comparable to the averaging techniques that have been used in previous autonomic conditioning studies it nonetheless reveals essentially the same conclusions as the previous analysis. Again Groups I and II show evidence of conditioning whereas Groups III-A and III-U do not.

An analysis of variance following

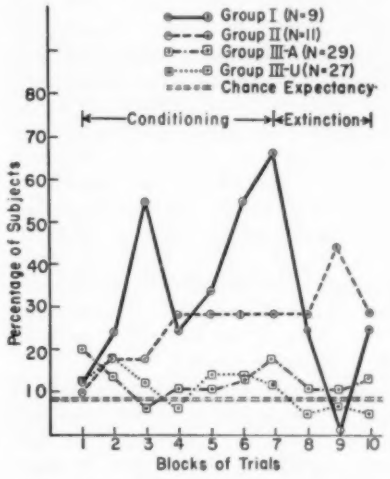


FIG. 1. Percentage of *Ss* whose maximum cardiac response difference follows the stimulus BOAT.

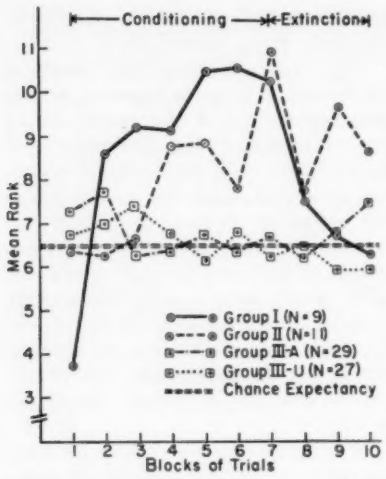


FIG. 2. Mean rank of cardiac response difference which follows the stimulus BOAT.

Lindquist's (1953) Plan I for the conditioning trials only was used to test the significance of the effects in Fig. 2. There was a significant Between-Group effect ($F = 6.35$, $df = 3/72$, $P < .01$) and a significant Group × Trial Blocks interaction ($F = 2.27$, $df = 18/432$, $P < .01$). The significant Between-Group effect was analyzed further by means of t tests using the appropriate error terms from the analysis of variance. On the seventh trial block, the difference between Groups I and II was not significant ($t = .14$, $df = 72$), nor was the difference between Groups III-A and III-U ($t = .55$, $df = 72$). However, both Groups I and II differed significantly from both Group III subgroupings ($t = 2.88$, $df = 72$, and $t = 2.95$, $df = 72$, respectively).

A separate but similar analysis of variance was applied to the extinction trials using the rank scoring. The only significant effect was Between Groups ($F = 3.32, df = 3/72, P < .05$). Further analysis by the t test revealed the significant Between-Group effect to be due to the differences between Groups II and III-U.

Cardiac responsiveness.—While the above results are quite clear in indicating the role of cognitive expectancies on the acquisition of conditioned heart rate responses, it is also informative to investigate the effects of these cognitive expectancies upon other characteristics of heart rate behavior, specifically heart rate responses to nonconditioned words. This was done by determining the cardiac response difference to all words except BOAT for each trial block for each *S*. In Fig. 3 the average response difference for these 11 words has been plotted as a function of trial blocks during conditioning and extinction for each of the four groups.

There is a tendency for the cardiac response difference to words other than the CS to decrease throughout the conditioning and extinction trials. Also there is a marked difference be-

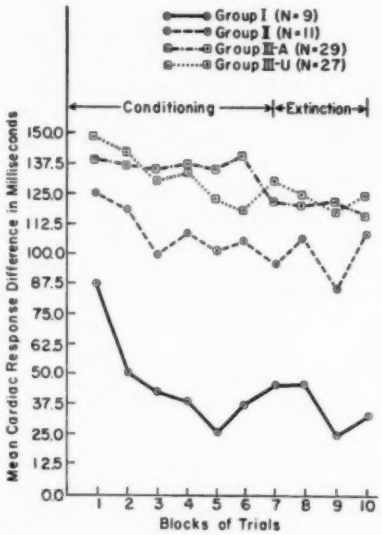


FIG. 3. Mean cardiac response difference following all stimulus words excluding BOAT.

tween groups. Group I shows less reactivity to nonconditioned words throughout the conditioning and extinction sessions with Group II falling between Group I and Groups III-A and III-U. The latter two groups are indistinguishable in performance.

The significance of the effects shown in Fig. 3 was evaluated by an analysis of variance using Lindquist's Plan I. In this analysis conditioning and extinction trials were included in the same analyses. The only significant effects were for Between-Trial Blocks ($F = 4.13$, $df = 9/648$) and Between Groups ($F = 10.03$, $df = 3/72$).

Since the biggest group difference in Fig. 3 is due to Group I, it was desired to determine whether the different cognitive expectancies between Group II and Groups III-A and III-U had an effect upon the cardiac response difference to the nonconditioned stimuli. To determine this, the above analysis was repeated using only the latter three groups. Again there was a significant Trial Block effect ($F = 3.90$, $df = 9/676$), but the Between-Group effect ($F = 1.41$, $df = 2/64$) was not significant.

The above analyses have suggested that heart rate conditioning and heart rate behavior are in part a function of *S*s' verbalized expectancies concerning shock. We can ask the reverse question as to whether differences in heart rate behavior will predict *S*s' verbalizations. For this question the data of Group III are available. The *S*s in this group were asked to report the number of shocks they thought they had received to each of the stimulus words. Our previous analysis would suggest that words having a large cardiac response difference should have a higher number of reported shocks than words with a low cardiac response difference. To test this possibility the word giving the greatest cardiac response difference on the seventh conditioning trial was selected for each *S* along with the word producing the smallest cardiac response difference. The number of reported shocks to these two words was determined for each *S* in Group III. The mean number of shocks reported to the word with the greatest cardiac response difference was 1.37 as compared with 1.03 for the word with the smallest difference. A t test for correlated scores gave a value of 1.71 significant at the .05 level for a one-tailed test.

Tests of generalization.—In assessing generalization only the cardiac responses on the first trial block in the extinction series were examined. As will be recalled, the stimulus words for extinction contained not only the CS but words semantically similar and dissimilar and colors that were the same and different from the CS.

Detailed analyses of the data from Group III failed to yield any significant or suggestive evidence of either semantic or color generalization. Similarly, separate analyses of the data from Groups I and II who had shown conditioning, gave no evidence of generalization. Since the previous analyses had suggested that cognitive expectancy was an essential correlate of heart rate responses, and these *S*s had not verbalized semantic or color relations, the lack of generalization might have been anticipated.

Discussion

Conditioning was evident in Groups I and II but here all *S*s had clearly verbalizable expectancies concerning the relationship between stimulus and shock. The heart rate response of *S*s in these two groups shows a further correspondence with their cognitive expectancies in the extinction behavior. Group I *S*s were told following the seventh conditioning

trial that there would be no further shocks. The data are clear in showing almost complete extinction on the first extinction trial. Group II *Ss* on the other hand were not informed of the termination of the conditioning trials and there is little or no evidence of extinction of their heart rate response during the three extinction trial blocks.

Group III *Ss* who had received a minimum amount of information concerning the relationship between the stimuli and shocks prior to the conditioning session show no evidence of having conditioned heart rate responses. Even when this group of *Ss* is subdivided into subgroups based upon their verbal expectancies of shocks obtained postexperimentally there is no evidence that the subgroup showing some verbal discrimination between the shock and the nonshock stimuli has conditioned. It may be that it was only in the last of the conditioning trials that these *Ss* began to form cognitive expectancies concerning the relationship between specific stimuli and shock. Thus there was insufficient time for conditioned heart rates to occur.

But even among the Group III *Ss* there is some evidence of a relation between cognitive expectation and heart rate behavior. When the stimulus word with the greatest heart rate response is compared with the smallest heart rate response on the last conditioning trial, it is found that these words are discriminated in the terms of *S*'s verbalized shock expectancies.

The effects of cognitive expectancies is also apparent on other aspects of heart rate behavior in the experimental situation. Group I *Ss* who knew that only one word would be shocked during the experimental session showed an appreciably smaller heart rate response to the nonconditioned stimulus words throughout the conditioning and extinction session. Similarly Group II *Ss* who knew that words other than the CS would be shocked only once during the session showed less heart rate response to these nonconditioned stimuli than did the Group III *Ss* who had no definite expectation of relationships between stimuli and shock.

The absence of any stimulus generalization along either the semantic or the color dimension is also consistent with the findings concerning the importance of cognitive expectancies on the occurrence of heart rate response changes to the stimuli. Since conditioning occurred only in the Group I and II *Ss* generalization can only be expected to occur in these groups. However, *Ss* in these groups had clearly verbalizable expectancies concerning the relationship between the conditioned stimulus and shock. Since their verbalizations during the inquiry period did not express relationships between colors or semantic class and the occurrence of shock, the hypothesis that cognitive expectancy determines heart rate behavior would predict a failure of generalization to occur among these *Ss*.

While the data considered so far are quite unequivocal in demonstrating a relationship between heart rate behavior and cognitive expectancy the question can be raised as to whether the cognitive expectancy is necessarily prior to or a determiner of the heart rate phenomena. There are several factors in the present experiment that would indicate that cognitive expectancy is a determiner of the heart rate response rather than something that develops concurrently with the conditioned heart rate. In previous studies the typical procedure has been to allow awareness or expectancies to develop along with the CR. In this type of design it is impossible to determine causal relationships. However, by manipulating the expectancies of our *Ss* prior to the beginning of the conditioning sessions we have largely controlled the expectancies that existed in our *Ss* before conditioning of the heart rate occurred.

There is further evidence of the primacy of cognitive expectancies in the extinction behavior of the Group I *Ss*. The knowledge that no further shocks would occur was sufficient to produce almost complete extinction without experiencing the CS in the absence of the

UCS. This result is consistent with that of Notterman et al. (1952).

The results we have obtained are also quite consistent with those of Branca (1957), who reports marked correspondence between conditioned autonomic behavior and *Ss*' verbalizable expectancies in the conditioning situation.

One further comment is in order concerning the low relationship between cognitive expectancy and heart rate behavior. While the above evidence shows a definite relationship between cognitive expectancies and heart rate behavior the relationship is certainly not a very high one. Eriksen (1958) has pointed out elsewhere the existence of large noncorrelated errors between different response systems such as the verbal and the autonomic and the present low relationships are probably a reflection of the amount of error in cognitive expectancies as well as heart rate behavior. In view of other evidence it is most likely that the largest source of error is in the heart rate response rather than in *Ss*' verbalization of his expectancies.

Summary

The conditioning and semantic and color generalization of the heart rate was studied as a function of different cognitive expectancies of the *Ss*. Cognitive expectancies were manipulated prior to conditioning by means of different instructions to the three experimental groups.

A high correspondence was found between heart rate and verbalizable expectancies. Clear evidence of heart rate conditioning was obtained only in those cases where *S* could verbalize the relationship between the CS and UCS. Those *Ss* who were informed that there would be no more shocks at the beginning of the extinction trials showed almost a complete loss of the CR without experiencing nonreinforced presentations of the CS. Heart rate responses to nonconditioned stimuli were also found to vary as a function of cognitive expectancies. There was some indication that observed differences in heart rate could be used to predict differences in verbalized expectancies.

There was no evidence of either semantic or color generalization of the conditioned heart rate. This finding was considered consistent with the above since *Ss* did not include such generalized expectancies in their verbalizations.

References

Branca, A. A. Semantic generalization at the level of the conditioning experiment. *Amer. J. Psychol.*, 1957, **70,** 541–549.

Chatterjee, B. B., & Eriksen, C. W. Conditioning and generalization as a function of awareness. *J. abnorm. soc. Psychol.*, 1960, **60,** 396–403.

Diven, K. Certain determinants of conditioning of anxiety reactions. *J. Psychol.*, 1937, **3,** 291–308.

Eriksen, C. W. Unconscious processes. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1958*. Lincoln: Univer. Nebraska Press, 1958. Pp. 169–227.

Haggard, E. A. Experimental studies in affective processes: I. Some effects of cognitive structure and active participation on certain autonomous reactions during and following stress. *J. exp. Psychol.*, 1943, **33,** 257–284.

Lacey, J. L., & Smith, R. L. Conditioning and generalization of unconscious anxiety. *Science*, 1954, **120,** 1045–1052.

Lindquist, E. F. *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin, 1953.

Notterman, J. M., Schoenfeld, W. N., & Bersch, P. J. Partial reinforcement and conditioned heart rate response in human subjects. *Science*, 1952, **115,** 77–79.

Razran, G. Stimulus generalization of conditioned responses. *Psychol. Bull.*, 1946, **46,** 337–365.

Thorndike, E. L., & Lorge, I. *Teacher's word book of 30,000 words*. New York: Teachers College, Columbia University, Bureau of Publications, 1944.

(Received August 5, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 280–287

TEST OF THE HYPOTHESIS OF PSYCHOLOGICAL REFRACTORY PERIOD[1]

JACK A. ADAMS

University of Illinois

Telford (1931) found that the second of two reaction time measures was lengthened when the spacing of two successive stimuli was reduced to 500 msec., and his hypothesis of psychological refractory period has motivated a number of confirming studies and analyses (Adams, 1961; Craik, 1948; Davis, 1956, 1957, 1959; Marill, 1957; Vince, 1948; Welford, 1952, 1959). A prominent interpretation of these findings is that an incoming stimulus is subjected to a central decision process before discharge occurs down the motor nerves, and a second stimulus impinging during this decision time is either disregarded, degraded, or delayed in immediate memory until the decision mechanism has been cleared. A general conclusion is that *S* is a one-channel data processing system.

Hick (1948), Poulton (1950), and Elithorn and Lawrence (1955) have suggested a counterexplanation in terms of expectancy. In this usage, "expectancy" refers to *S* having learned certain properties of the statistically defined time relationships between the first and the second stimulus presented over a relatively long series, and *S* is thought to be most alert for responding when the interstimulus interval is somewhere in the vicinity of the mean delay. When a very short interval occurs, *S* is not ready, and his response to the second stimulus is lengthened. An implication of the expectancy hypothesis is that *S* can be a multichannel system when conditions allow the acquisition of appropriate expectancies.

All findings can be explained about equally well by both hypotheses, but a discriminating test seems possible by manipulating the statistical structure of interstimulus time intervals. So far, studies of this topic have used only highly uncertain, random interval distributions (low redundancy). The expectancy hypothesis would predict that refractoriness is a function of the statistical structure of time intervals by making *S* more expectant for certain classes of intervals, while the one-channel hypothesis would regard refractoriness a function of intervals but not statistical structure. The investigation reported here manipulated the redundancy of interstimulus intervals in a two-dimensional, bisensory discrete tracking task where an audio signal occurred with or lagged a visual signal by a defined time interval.

METHOD

Apparatus.—The discrete tracking apparatus was used (Adams & Chambers, 1962). This device can be used as a two-dimensional bisensory discrete tracking task where the visual and the audio inputs each have three discrete states, or as a one-dimensional tracking task with either audio or visual stimuli. In the visual dimension, *S* had three horizontally arranged jeweled stimulus lights (red, white, and green) in front of him at eye level, and these lights came on in a repetitive sequence at defined time intervals (see below). Beneath each stimulus light was a small neon

[1] This research was supported by the United States Air Force under Contract No. AF 49(638)-371, monitored by the Air Force Office of Scientific Research of the Air Research and Development Command. Acknowledgment is due the assistance of William F. McDonald and Luther J. Tromater.

response feedback light that informed *S* of the position of his control, and this direct display of response feedback cues, as well as stimulus lights, made the task one of pursuit tracking. The *S* had to keep the feedback light aligned with the frequently changing stimulus light as much as possible. The audio dimension had a 600-, 800-, and 1000-cycle pure tone as stimuli, and *S* heard them over a headset and responded with the same type of control as used for the visual dimension. The auditory error coding was in terms of a complex tone. When *S* had the control in the correct position, he heard a pure tone but, when he was wrong, he heard a complex tone made up of two of the three fundamental frequencies. The correct stimulus was presented as a pure tone in the complex, and, superimposed on it was a second tone, rapidly interrupted, whose frequency was determined by the position of the control. Thus, when *S* was in error he had feedback on the correct stimulus and the wrong control position, and this was formally equivalent to the pursuit tracking format of the visual task. Moreover, being pursuit tracking, *S* had direct feedback of his response time to stimulus change in each dimension. He could see how long it took for the visual feedback light to become aligned in the correct position, and similarly he could hear the duration of the interrupted tone. The instructions were to always respond and nullify error as quickly as possible and, in the bisensory task, to give equal attention to both dimensions.

Response was with a 2-in. control stick mounted on the wide armrest of *S*'s chair, and he could have a control for one or both hands, depending on whether tracking was unisensory or bisensory. The control moved freely through its arc, although electrically it could only assume three states. Audio or visual stimuli could be switched to either control. The direction of control movement was always horizontal. An Esterline-Angus operations recorder allowed *E* to completely record all stimulus and response events and their time relationships on a trial.

Procedure.—There were three groups of *Ss*, distinguished by the statistical distribution of time delay intervals governing the amount that an audio stimulus followed a visual one in two-dimensional bisensory tracking. Each group was given three practice sessions on different days, usually within the same week. A session was 12 2-min. trials, with 5 min. rest between Trials 4 and 5, and 8 and 9, and 50 sec. intertrial rest for the remainder of the trials. The basic phenomenon of psychological refractory period should be manifest as delay in response to the audio stimulus when it follows the visual stimulus too closely, and it is necessary to have a unisensory audio control condition to demonstrate this delay. In addition, it was considered advisable to have a unisensory visual control condition to see if time uncertainty influenced response to the visual stimulus as well as audio. Each *S* provided his own unisensory control measures. For the 12 trials of a session, 4 were unisensory visual, 4 unisensory audio, and 4 bisensory. The order of these three task conditions was counterbalanced among *Ss* of each group, and the particular order assigned to an *S* was the same on each session. Also, within each group, the assignment of visual and audio signals to left and right hands was counterbalanced. The operations recorder was used on Trials 2 and 3 of each block of 4 trials in Session 3 to provide a detailed analysis of individual responses.

Stimulus series.—The type, duration, and interstimulus intervals of audio and visual events were programed on punched tape and automatically read by a motorized tape reader. A single stimulus input tape was constructed for each group, and it had 60 audio-visual bisensory stimulus pairs on each trial. A given group used the same tape on each session. Whenever a unisensory trial was required, *E* disconnected the unwanted stimulus series in the other sensory dimension from the presentation. The 60 visual events on a trial had durations of 1.5, 2.0, and 2.5 sec., and there were 20 of each duration. The order of the 60 visual events was separately randomized for each trial. All groups had the same order of visual events on a given trial.

The audio delay intervals were in the range where the phenomenon of psychological refractory period was expected to be maximal, as well as somewhat beyond this range, and were 0, 100, 200, 400, and 800 msec. The approach was mainly to manipulate the frequency of audio delay intervals of 100 msec., where the phenomenon of psychological refractory period is known to be high. If expectancy is a significant explanatory mechanism for behavior, the amount of degradation in audio RT should be influenced by the frequency of events at the very small delay intervals. The *S* should be more expectant with a greater frequency of small intervals and consequently should have less decrement in audio RT. On the other hand, if expectancy is not a relevant explanatory framework, no difference should be expected. The statistical distribution of audio time delay intervals on a trial for each group, mean delay, and percent-

TABLE 1

NUMBER AND PROBABILITY OF DELAY INTERVALS BETWEEN VISUAL AND AUDIO STIMULUS ON A BISENSORY TRIAL FOR EACH GROUP

Audio Delay Interval (in Msec.)	Group					
	LU		MU		HU	
	p	N	p	N	p	N
0	.05	3	.10	6	.20	12
100	.80	48	.60	36	.20	12
200	.05	3	.10	6	.20	12
400	.05	3	.10	6	.20	12
800	.05	3	.10	6	.20	12
Mean (msec.)	150		200		300	
% Redundancy	52		24		0	

age redundancy (Attneave, 1959) are shown in Table 1. The groups were designated low uncertainty (LU), medium uncertainty (MU), and high uncertainty (HU). The order of the delay intervals was separately randomized for each trial.

It should be emphasized that the experiment deals only with time uncertainty in tracking, not event uncertainty as investigated by Adams and Chambers (1962) with this task. Each sensory dimension was a repetitive series of the three stimulus events which simply required *S* to move the control back and forth. And, on the bisensory trials, the same audio and visual events were always paired together. Thus, the stimulus series always had *event certainty*. However, because the time patterning of events was statistically determined, the stimuli always had *time uncertainty*, which was the basic experimental variable.

Subjects.—There were 18 *S*s in a group. The 54 *S*s were university male undergraduates who were paid for their participation. They were randomly assigned to groups.

Performance measurement.—Overall proficiency in discrete tracking, such as measured by time on target, is a function of (*a*) off-target time between the onset of the stimulus and the onset of the response, whether the response is correct or not, (*b*) number of errors, i.e., movements of the control to the wrong position, and (*c*) duration of each error before it is corrected. Because the motor movements required were simple and repetitive, errors were negligible and so proficiency primarily was determined by the off-target time of a correct response to change of a stimulus light or tone. These time values for individual responses are called response times (RT), and are distinguished from classical reaction time where special steps are taken to see that a response is nonanticipatory and always follows the stimulus (Woodworth, 1938). In our discrete tracking task a response could follow a stimulus as in classical reaction time studies, or it could be anticipatory as might be expected from a time-sensitive *S* who had acquired expectancy states. A better understanding of expectancy was hoped for by freely allowing, measuring, and analyzing anticipatory responses. The RTs were measured as the difference in milliseconds, between the onset of a stimulus and the occurrence of a correct response to it. Consistent with Poulton (1952), Adams and Xhignesse (1960), and Adams and Chambers (1962), a positive sign was assigned when the response preceded the stimulus, and a negative sign when it followed.

The basic analysis was conducted on an RT score for an *S*, which was the algebraic mean of all his response times to individual stimuli of a particular set of stimulus events. Each *S* had a Unisensory Visual RT, a Unisensory Audio RT, and, for each audio delay interval, a Bisensory Visual RT and a Bisensory Audio RT score.

RESULTS

Unisensory-bisensory audio comparisons.—Figure 1 is a plot of group

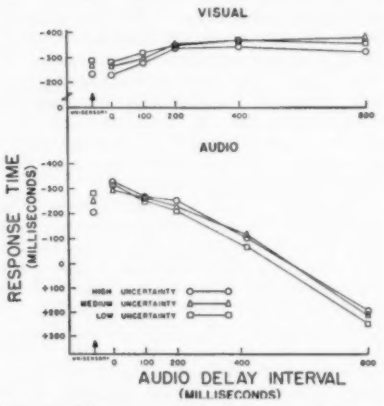


FIG. 1. Mean response times to unisensory and bisensory visual and audio stimuli.

means for Unisensory and Bisensory RT scores for visual and audio. The lower part of Fig. 1 presents the data that are most critical for assessing the one-channel hypothesis. The hypothesis predicts that Bisensory Audio RT will be lengthened when the interstimulus interval is very brief because *S* must have a finite period of time to process the visual stimulus and its response. Figure 1 shows the expected effect, and it is evident for all groups, particularly at the zero audio delay interval. Using the *t* test for related measures, a comparison was made for each group between Unisensory Audio RT scores and Bisensory Audio RT scores at delay intervals of 0, 100, and 200 msec., where the effects of psychological refractory period should be most evident. For all three groups, the Bisensory Audio RT was significantly poorer ($P < .01$) than the Unisensory Audio RT at the zero interval, but only Group HU had significant retardation at the 100- ($P < .01$) and 200- ($.01 < P < .05$) msec. audio delays. Beyond the three initial points, the Bisensory Audio RTs had near zero or positive values and indicate the presence of anticipatory responding.

The greater persistence of significant decrement in the Bisensory Audio RT scores for Group HU is in line with the expectancy hypothesis. However, an analysis of variance of Unisensory Audio RT scores gave a significant *F* ratio ($.01 < P < .05$), with Group HU having the lowest score, which could accentuate the differences. Examination of the unisensory audio data suggested that this was due to the differential presence of anticipatory responding among the groups. Because virtually all studies of psychological refractory period have used discrete reaction time tasks where special steps are taken to avoid the influences of anticipation in performance measures, it is of special interest to see if the same support can be found for the expectancy hypothesis when only nonanticipatory measures are used. Using the same approach as related experiments that dealt with anticipation (Adams & Chambers, 1962; Adams & Xhignesse, 1960), three classifications of individual RT measures were made as a means of examining anticipatory behavior: Beneficially Anticipatory RT values in the range of ±133 msec. that had less off-target time than ideal RT values (Klemmer, 1956, 1957) and, by giving little or no off-target time, allow the reasonable inference that positive anticipatory mechanisms were operating; Nonanticipatory RT values which were less than −133 msec., and in the range for classical reaction time where *S*'s response occurs substantially after the stimulus; and Detrimentally Anticipatory RT values which were greater than +133 msec. where *S* responded well ahead of the stimulus and could net as much, and often more, off-target time than if he had waited for the stimulus to occur and responded as in classical reaction time. An *S*'s RT score in each of these classifications was the algebraic mean of all his individual RTs in a classification for a particular experimental condition.

Groups LU, MU, and HU had 3, 4, and 11% of their individual unisensory audio RTs, respectively, that could be classified as either beneficially or detrimentally anticipatory. Using only Nonanticipatory Audio RTs, a Nonanticipatory Audio RT score was computed for unisensory audio and for bisensory audio under delay conditions of 0, 100, and 200 msec. A *t* test for related measures was made between unisensory and bisensory audio under each of the three delay

conditions. None of the tests achieved the .05 level of significance for Group LU, and Group MU had a t ratio at the .05 level for the zero interval only. Group HU had a t ratio significant at better than the .01 level for the zero delay, and t's significant at the .05 level for both the 100- and 200-msec. intervals. The same trend in support of the expectancy hypothesis is evident as before.

Unisensory-bisensory visual comparisons.—The upper part of Fig. 1 shows group mean RTs for unisensory and bisensory visual performance. An analysis of variance test revealed unisensory visual performance to be significantly poorer than unisensory audio ($P < .01$). While this is the traditional finding of visual RT being slower than audio RT, the two unisensory series are not directly comparable because the unisensory visual series had less time uncertainty than unisensory audio. The dominant trend in bisensory is for *S*s to temporarily withhold their visual response as a function of the audio delay interval, and this trend is about the same for all groups. As with audio performance, t tests for related measures were run between unisensory and bisensory visual RT scores. None were significant at the zero delay interval but all were significant at 100 msec. and beyond ($P \leq .05$).

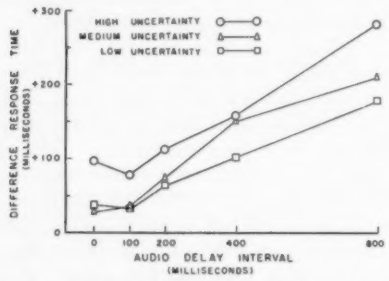


FIG. 2. Mean Difference RT values showing time lag between the visual and the audio response in bisensory tracking as a function of the audio delay interval.

Characteristics of bisensory performance.—To see if further insight could be obtained into reasons for the patterns of response decrement, a new measure was devised for the bisensory data, called the Difference RT, which gave the time between the two bisensory responses. This measure was used to check whether the amount and trend of decrement could be related to the way in which *S*s lagged the audio response behind the visual response as a function of uncertainty of the series. The expectancy hypothesis suggests that Group HU would lag the most because they would have developed higher expectancies for longer delay intervals. The Difference RT is defined by the formula (Bisensory Visual RT) − (Bisensory Audio RT) + (Audio Delay Interval), where the value for the delay interval is always positive, and the algebraic convention for RT is retained. The Difference RT is positive when the audio response follows the visual response, and negative when it precedes. Figure 2 shows the plot of mean Difference RTs. All individual RT values, plus and minus, entered the group means in Fig. 2. The three groups all have the same general trend, with Group HU at a higher level throughout (longer lags). Figure 2 shows that when the audio delay interval was zero or quite small, *S*s lagged the audio stimulus by a small amount. But as the audio delay interval increased, *S*s increased the lag of audio responses correspondingly. It would appear that the simultaneous occurrence of the visual and the audio stimulus is a cue for a rapid visual response and a small lag of the audio response. But if there is audio delay, and the visual stimulus

TABLE 2

PERCENTAGES OF THE THREE POSSIBLE ORDERS OF BISENSORY RESPONSE PAIRS

Order of Response Pair	Group and Audio Delay Interval (in Msec.)																	
	Low Uncertainty (LU)						Medium Uncertainty (MU)						High Uncertainty (HU)					
	0	100	200	400	800	All Intervals	0	100	200	400	800	All Intervals	0	100	200	400	800	All Intervals
Visual first	36	30	37	40	44	32	32	34	40	55	41	37	56	53	60	57	61	57
Audio first	25	15	14	20	16	16	28	18	24	16	20	20	11	13	12	12	10	12
Simultaneous	39	55	49	40	40	52	40	48	36	29	39	43	33	34	28	31	29	31

Note.—Percentages are based on the total number of correct response pairs made by a group under the conditions specified by a column heading.

is on momentarily by itself, then these delay states of affairs become a cue for delays in the visual response and longer lags of the audio response, although Fig. 1 shows that when the delay interval is 800 msec. the lag is poorly timed because Audio RTs are detrimentally anticipatory by 200–300 msec. Thus, *S* learns to interpret the immediate temporal properties of stimuli, and responds differentially to them, and Fig. 2 shows that the overall probability structure of delay intervals tends to differentially influence group mean performance in each case. The small intervals of 0, 100, and 200 msec. are most critical for issues in question, and an analysis of variance (Lindquist, 1953, Type I) was performed on the Difference RTs, with uncertainty a between-*S*s variable and audio delay a within-*S*s variable. Both main effects were significant at the .02 level. The same analysis was performed on Difference RTs computed from nonanticipatory responses, and again both main effects were significant ($P < .01$).

The Difference RTs show that Group HU lags the audio response more, which means longer RTs when audio delay intervals were small, and thus the decremental effect which is the evidence for the one-channel hypothesis. A related implication is that Group HU should have more bisensory response pairs occur in a visual-audio sequence rather than in a joint, simultaneous fashion. To evaluate this, the three possible sequences of response pairs were tabulated for each group, and the results are shown in Table 2. The order of responses can be either visual first, audio first, or simultaneous responding. The criterion for simultaneity was the two responses within ± 33 msec. (Adams & Chambers, 1962). It is noteworthy that each audio delay interval has a percentage of response pairs in each of three orders, and there is no marked tendency towards "grouping," or simultaneous responding, to be particularly concentrated at very small intervals as Vince (1948) and Welford (1959) hypothesize would occur because of a stimulus pair being perceived as an entity under these conditions. Group differences are evident, particularly when the percentage is taken over all intervals, and the trends are ordered in accord with the expectancy view. Group HU had a dominant tendency to make the visual response first and lag the audio response. Group LU had stimulus pairs that almost always occurred with a very small time separation, and

they tended to respond with more simultaneity. Group MU had an intermediate position. To test these differences, each *S* was given a score in each of the three categories of response order that was the total number of responses made over all delay intervals. These were called the Visual First score, the Audio First score, and the Simultaneous score. A simple analysis of variance was performed for each of the three sets of scores, and the Visual First scores and the Simultaneous scores had between-groups differences that were significant at the .01 level. Audio First scores were significant between the .01 and .05 levels. The importance of Audio First pairings is not readily interpretable, but it could represent a tendency towards error in closely timed simultaneous responding.

Discussion

The results are consistent with the expectancy hypothesis. Decrement in response to the second of two closely spaced stimuli was greatest when the stimulus series had high time uncertainty, and was reliably less when time uncertainty was moderate or low. The reason was that *S*s in the high uncertainty group had a greater likelihood of receiving a visual-audio stimulus sequence with a relatively long audio delay, and they learned to respond more frequently with a visual-audio order and to lag the audio response longer. The *S*s had an expectation, or set, for the visual stimulus to come on first and for the audio stimulus ordinarily to be delayed and require a lagged response. Even when the two stimuli were presented simultaneously and *S*s had direct and immediate information that a lag was not required, the audio response was still delayed. In fact, the Bisensory Audio RTs at the zero interval were the longest of any obtained. Thus, in this study, the decremental effect that has come to exemplify psychological refractory period emerges as a learned tendency to respond with a visual-audio sequence and to lag the audio response when extensive practice has been given under conditions of temporal uncertainty.

The findings do not allow unequivocal rejection of the one-channel hypothesis, because even Group LU with low temporal uncertainty in their stimulus series had a significant amount of decrement in Bisensory Audio RT scores for stimulus pairs that occurred simultaneously. Remembering the wealth of temporal anticipation effects present in the data, this finding could be interpreted to mean that there is a one-channel decision mechanism that momentarily delays each S-R sequence, but that it cannot find a meaningful place in any scientific account of human behavior that does not give central focus to temporal expectancy states. Nevertheless, when these data are weighed with those of the Adams and Chambers (1962) study, there are defensible grounds for questioning the one-channel view. Using the same task, Adams and Chambers presented findings that dovetail with those here because one of their experimental conditions was complete temporal and event certainty. Not only did they find impairment absent, but they found that time and event certainty were the conditions for bisensory performance actually being superior to the aggregate unisensory control performances because of the influences of temporal anticipation. The conclusion from these two investigations is that given sufficient temporal certainty of events, *S* can process at least two simultaneous stimulus series with the same proficiency as a single stimulus series providing there is event certainty. It remains to be fully determined, however, whether a necessary central delay time exists for the resolution of *event uncertainty*. Adams and Chambers found first evidence for impairment under conditions of event uncertainty and temporal certainty.

Summary

An experiment was performed to test the hypothesis of psychological refractory period

that is offered to account for the established finding that response to the second of two closely spaced stimuli shows decrement. One line of explanation argues for a central decision time, where time must be allowed for processing the first stimulus and response before the second sequence can be undertaken. A competing explanation is the expectancy hypothesis which ascribes decrement to *S*'s past experience with the random array of interstimulus intervals that is usually used in experiments on this topic. Through practice, *S* comes to expect a longer delay and the decrement is because he is not optimally ready to respond.

The experiment involved a two-dimensional, bisensory discrete tracking task. The statistical structure of interstimulus time intervals was the experimental variable aimed towards discriminating between the two hypotheses by asking if decrement could be a function of the temporal organization of stimuli. The results supported the expectancy hypothesis. Reliably less decrement was found for *S*s who trained on a stimulus series with a predominance of small time intervals and could learn behavior appropriate to them.

References

ADAMS, J. A. Human tracking behavior. *Psychol. Bull.*, 1961, **58,** 55–79.

ADAMS, J. A., & CHAMBERS, R. W. Response to simultaneous stimulation of two sense modalities. *J. exp. Psychol.*, 1962, **63,** 198–206.

ADAMS, J. A., & XHIGNESSE, L. V. Some determinants of two-dimensional visual tracking behavior. *J. exp. Psychol.*, 1960, **60,** 391–403.

ATTNEAVE, F. *Applications of information theory to psychology.* New York: Holt, 1959.

CRAIK, K. W. J. Theory of the human operator in control systems: II. Man as an element in a control system. *Brit. J. Psychol.*, 1948, **38,** 142–148.

DAVIS, R. The limits of the "psychological refractory period." *Quart. J. exp. Psychol.*, 1956, **8,** 24–38.

DAVIS, R. The human operator as a single channel information system. *Quart. J. exp. Psychol.*, 1957, **9,** 119–129.

DAVIS, R. The role of "attention" in the psychological refractory period. *Quart. J. exp. Psychol.*, 1959, **11,** 211–220.

ELITHORN, A., & LAWRENCE, C. Central inhibition: Some refractory observations. *Quart. J. exp. Psychol.*, 1955, **7,** 116–127.

HICK, W. E. The discontinuous functioning of the human operator in pursuit tasks. *Quart. J. exp. Psychol.*, 1948, **1,** 36–44.

KLEMMER, E. T. Time uncertainty in simple reaction time. *J. exp. Psychol.*, 1956, **51,** 179–184.

KLEMMER, E. T. Simple reaction time as a function of time uncertainty. *J. exp. Psychol.*, 1957, **54,** 195–200.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education.* New York: Houghton Mifflin, 1953.

MARILL, T. The psychological refractory phase. *Brit. J. Psychol.*, 1957, **48,** 93–97.

POULTON, E. C. Perceptual anticipation and reaction time. *Quart. J. exp. Psychol.*, 1950, **2,** 99–112.

POULTON, E. C. Perceptual anticipation in tracking with two-pointer and one-pointer displays. *Brit. J. Psychol.*, 1952, **43,** 222–229.

TELFORD, C. W. Refractory phase of voluntary and associative responses. *J. exp. Psychol.*, 1931, **14,** 1–35.

VINCE, M. A. The intermittency of control movements and the psychological refractory period. *Brit. J. Psychol.*, 1948, **38,** 149–157.

WELFORD, A. T. The "psychological refractory period" and the timing of high-speed performance: A review and a theory. *Brit. J. Psychol.*, 1952, **43,** 2–19.

WELFORD, A. T. Evidence of a single-channel decision mechanism limiting performance in a serial reaction task. *Quart. J. exp. Psychol.*, 1959, **11,** 193–210.

WOODWORTH, R. S. *Experimental psychology.* New York: Holt, 1938.

(Received August 14, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 288–294

THE COURSE OF EMOTIONALITY IN THE DEVELOPMENT OF AVOIDANCE [1]

HOWARD S. HOFFMAN AND MORTON FLESHLER

Pennsylvania State University

Dual process theories of avoidance (Hull, 1943; Miller, 1951; Mowrer, 1950; Solomon & Wynn, 1953) are predicated on the assumption that the acquisition of the instrumental avoidance response involves, and moreover depends upon, the concurrent acquisition of a conditioned emotional response (sometimes identified as fear or anxiety). The purpose of the present investigation was to assess the conditioned emotional responses (CERs) that occur during the acquisition of avoidance.

The results of such a study provide a test of dual process theory because the failure to demonstrate a meaningful relationship between emotional and instrumental behavior would cast considerable doubt on the dual process position. To the extent that meaningful relationships are observed, however, the results would provide an opportunity to evaluate the sequence of changing interactions between instrumental and emotional behaviors which, according to dual process theory, occur during acquisition. Theoretical specification of this sequence of interactions has, of necessity, been somewhat speculative. Experiments which have documented the acquisition of avoidance behavior have either inferred the state of concurrent emotionality from crude indices (bolus counts, incidence of freezing, trembling, etc.) or more frequently from the effects of various experimental operations upon the avoidance response itself (Solomon & Brush, 1956). Studies which have focused upon the development of emotional responses, on the other hand, have seldom employed avoidance procedures. One exception, however, is a study by Black (1956) which assessed cardiac reaction during the acquisition of avoidance. Although Black's results appear to offer some support for dual process theory, they are difficult to interpret because he also found that about two-thirds of the total cardiac reaction could be attributed to the muscular occurrence of the instrumental response itself.

The approach of the present study was to track the course of emotionality by developing an avoidance response while *S*s were engaged in positively reinforced ongoing behavior. In this paradigm, the index of emotionality is the decrement in rate of positively reinforced responding (conditioned suppression) which occurs during the presentation of a warning signal which precedes electrical shock. Conditioned suppression has been examined extensively and found to provide a sensitive and reliable index of emotional responses (Brady & Hunt, 1955; Estes & Skinner, 1941). The present study differs from the usual conditioned suppression experiment in one major respect. In studies of conditioned suppression, the noxious event is unavoidable. In the present arrangement, on the other hand, a specific instrumental response, during

[1] This research was supported by National Institute of Mental Health Grant M-2433, and also by a grant from the Pennsylvania State University Central Fund for Research. The authors wish to thank Peter Day and Edmund Sequin for their assistance in the course of this study.

the CS, terminates that stimulus and prevents the noxious event.

METHOD

Subjects.—The *S*s were 12 female rats of Sprague-Dawley stock. They were approximately 10 mo. old at the start of the experiment.

Apparatus.—The experimental chamber was a sound insulated Skinner box fitted with a one-way vision observation port. In the middle of the front wall was a recess to which Noyes food pellets were delivered. On either side of the recess was a manipulandum. To the right was a bar of 0.25 × 0.75 in. aluminum which projected 1.5 in. into the test chamber at a height of 1.5 in. and which actuated a microswitch when a downward force of 20 gm. was applied. At the left, 1.5 in. above the floor was a plate (1.5 in. square) which protruded 0.5 in. from, and was parallel to, the front wall. When a horizontal force of at least 15 gm. was applied perpendicular to the plate, it actuated a second microswitch.

The walls, manipulanda, and grid floor were wired to carry electrical shock, and during shock the polarity of the grid bars, the manipulanda, and the walls was continuously scrambled, so as to make unauthorized escape highly improbable. The shock power was supplied by an Applegate constant current stimulator set at 1.5 ma.

Acoustic signals were delivered through a 5-in. speaker mounted on the back wall of the chamber. Out of a second speaker, mounted at the side, white noise was continuously presented in order to mask sounds having external origin. The warning signal was a pure tone at 3500 cps with an intensity of 88 db. re .0002 dynes per cm.^{2}, when measured in front of the speaker. The tone was generated by a Hewlett-Packard audio oscillator.

A series of timers, steppers, and relays was used to establish the several stimulus-response contingencies which the research demanded. The circuitry was such that a response (either bar press or plate press) was defined as the initial closure of the corresponding microswitch. Thus, holding responses had no effect on the program.

All stimuli and responses were recorded on an Esterline-Angus operations recorder. In addition, counters were used to record the number of responses occurring during the tone and during the periods which preceded and followed each tone. A Standard Electric timer was used to measure the latency of the bar press response to the tone.

Procedure.—All *S*s were treated alike in each of the several stages of training. The rats were first taught to escape from shock. After 20 min. of adaptation to the box, shock was turned on periodically; only a bar press terminated shock. After 50 presentations of shock, the median escape latency had stabilized at 0.75 sec. Twenty-five additional shocks were then delivered on each of two successive sessions in order to establish the bar press to terminate shock, as a well-learned habit. During these two sessions, the median escape latencies were 0.74 and 0.75 sec., respectively. The *S*s were then placed on restricted feeding and from then on were maintained at 75% of their previously determined free feeding body weights. During the period of weight reduction *S*s were placed in the box with the bar removed, so that generalized emotional responses could extinguish. In successive sessions, *S*s were trained to the food magazine, shaped to press the plate and were run on a six-response fixed ratio schedule of reinforcement. (The *S*s were exposed to the fixed ratio schedule in order to produce efficient response topographies.) A variable interval schedule of reinforcement with a mean of 30 sec. was in effect for the remainder of the experiment. Twenty sessions of plate pressing for food on the VI schedule were given to permit response rates to stabilize. By the end of these sessions, the rates had leveled off at a median value of 42 responses per min. These and all subsequent sessions each lasted 2.5 hr. and were run every other day.

During the next two sessions, while *S*s pressed the plate for food, 20 tones, each lasting 60 sec., were presented without shock at intervals of 5 to 7 min. This was done so that in later phases of the study, it would be possible to determine whether or not observed decrements in plate pressing during tone were attributable to the pairing of tone and shock.

The procedure was the same for the next three sessions except that the bar had been replaced. In these sessions a bar press which occurred during a tone terminated that tone. This procedural detail was initiated in order to determine whether or not, prior to its pairing with shock, the offset of tone would reinforce the bar press.

In the following session, avoidance training was initiated. The tone was programed to remain on for 70 sec. with shock programed to occur during its final 10 sec. A bar press at any time during the tone, but prior to

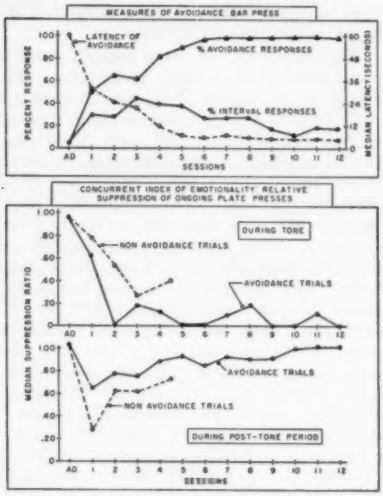


FIG. 1. Measures of avoidance and concurrent indices of emotionality throughout the course of acquisition. (AD refers to the final session of tone adaptation.)

shock onset, terminated that tone and prevented the occurrence of the shock. A bar press during shock terminated both the tone and the shock. Since the program of VI food reinforcement for plate presses was independent of the sequence of tone-shock pairings, food reinforcement could occur at any time during either the tone or the shock. Avoidance training was terminated after 12 sessions (20 tones per session).

Experimental measures.—For each *S* on each trial, the following information was obtained: (*a*) whether or not shock was avoided; (*b*) whether or not a bar press occurred in the 60-sec. interval that ended with the onset of tone (hereafter this interval will be identified as the pretone period); (*c*) the latency of the bar press during tone; (*d*) the rate of plate press during the 60-sec. pretone period, i.e., the number of plate presses during the pretone period divided by 60; (*e*) the rate of plate press during the tone, i.e., the number of plate presses during the tone divided by the duration of the tone (as determined by the latency of the bar press); and (*f*) the rate of plate press during the 60-sec. posttone period that began with offset of tone.

On each trial and for each *S*, Measures *d* and *e* and Measures *d* and *f* were then combined to form two suppression ratios. The first suppression ratio was the rate of plate press during tone divided by the rate of plate press during the pretone period, and serves as an index of the relative emotionality during the tone. The second suppression ratio was the rate of plate press during the posttone period divided by the rate of plate press during the pretone period, and serves as an index of the relative emotionality during the posttone period. These ratios were the entries employed in all subsequent statistical analyses of concurrent emotionality.

RESULTS

Figure 1 shows the session-by-session growth in avoidance, and also the changes in relative emotionality which develop concurrently. The top section of this figure shows the percentage of shocks that were avoided on each session, the percentage of times one or more bar presses occurred in the pretone period, and the median latency of the bar press during tone.

As seen in Fig. 1, during the final session of tone adaptation, ("AD" in the figure) *S*s seldom pressed the bar during either the tone or the pretone period. Since three sessions of tone adaptation had preceded this one, it is clear that, prior to its pairing with shock, the tone exhibited little, if any, control over the bar press. This result suggests that the tone itself was not intrinsically aversive, and that tone offset was not intrinsically reinforcing. Avoidance conditioning was instituted on Session 1. From Session 1 on, if the bar was pressed during the 60-sec. warning period, it terminated the tone and permitted *S* to avoid shock. If, however, a bar press during tone had a latency of more than 60 sec., it also occurred during shock and hence represented an escape response. Although both the shock and the tone were programed to terminate after 10 sec., no *S* ever permitted shock to remain on for longer than 2 sec. During the first session of avoidance conditioning the

median latency of escape was 0.79 sec. and the latency did not systematically change as sessions progressed. This result is in no way surprising since *S*s had previously received extended training on escape.

As seen in Fig. 1, the tendency to avoid increased with each session until by Session 6, 99% of the shocks were avoided. The tendency to bar press in the pretone period (interval responses) increased during the first three sessions, but with continued training gradually declined. The latency of the avoidance response simultaneously decreased, until by the end of Session 6, it had stabilized at about 6 sec.

A series of *t* tests for related measures was conducted on the frequencies of interval vs. avoidance responses during each of the first four sessions. The values of t ($df = 11$ for each) were 2.87, 4.01, 2.43, and 3.72 for Sessions 1 through 4, respectively. Since each of these values is significant at the .05 level (for a two-tailed test), it is clear that the three functions which appear in the top section of Fig. 1 represent the development of a well-discriminated avoidance behavior.

The bottom sections of Fig. 1 show the several indices of emotionality that were obtained on each session during the acquisition. The solid line in the middle section of Fig. 1 shows the median suppression ratio during tone for trials on which an avoidance response occurred. It can be seen that the tone caused essentially no suppression during the final adaptation session but that with the introduction of shock, it rapidly developed the capacity to suppress ongoing plate presses.

The dashed line in the middle section of Fig. 1 shows the median suppression ratio, during tone, on those trials during which *S* failed to avoid. Since, as sessions progressed, the number of these nonavoidance trials decreased rapidly, the data for Sessions 4 and 5 have been combined. No data are shown beyond Session 5 because, in this period, the number of nonavoidance trials was too small to yield reliable measures of suppression.

The suppression ratios obtained on avoidance trials provide an initial test of the dual process position, since a theoretical interpretation of the discriminated avoidance which occurred from Session 1 on (Fig. 1) must assume that at least during acquisition, a CER to the warning stimulus also occurred.

If, on a given trial, there were no systematic differences between the rate of plate press in the pretone period and the rate during tone, then during a given session, the number of suppression ratios above one should equal the number below one, i.e., the median suppression ratio would be one. If, as predicted by theory, the tone consistently evoked a CER, the median suppression ratio should be less than one. Sign tests conducted on the suppression ratios during tone, on avoidance trials, provided support for a dual process interpretation since each value of z, for Sessions 1 through 12, was greater than 4.0 ($P < .01$ for a two-tailed test, in each case).

There is, however, a question of the degree to which the low value of these ratios reflects the cessation in positively reinforced responding which must occur when *S* leaves the plate and executes the bar press. Two sources of information were used to assess this question. First, a series of sign tests were conducted on the suppression ratios, during tone, for nonavoidance trials. In Sessions 1, 2, and 3, as well as in Sessions 4 and 5 combined, the values of z were all greater than 2.82 ($P < .01$ in each

case). Thus, even on trials which were unconfounded by the occurrence of the avoidance response, the tone generated a substantial degree of suppression. Secondly, *S*s were observed throughout acquisition. In general, behavior in the presence of the tone appeared to involve considerable emotionality. Such plate presses as occurred, were performed in a tentative manner and the instrumental avoidance response involved an extremely slow sequence of movements, even on trials with latency as short as 6 sec. Both behaviors stood in sharp contrast to the quick energetic movements which typified the plate press during the pretone period and the instrumental bar press during shock. Finally, it may be noted that even if the suppression ratios on avoidance trials were adjusted for the maximum time *necessary* to execute the bar press (approximately .79 sec. as estimated from the latency of the bar press during shock), only a small increase would occur and the general configuration of the data would be unaltered.

A second assertion derived from dual process theory, is that during acquisition the tendency to avoid will be directly related to the CER magnitude. As seen in Fig. 1, suppression, during tone, was consistently greater on avoidance trials than on nonavoidance trials.

A Mann-Whitney test was used to assess the reliability of these data. For Sessions 1 and 2, the differences seen in the middle section of Fig. 2 were highly significant: $z = 3.15$ in Session 1 and $z = 4.02$ in Session 2 ($P < .01$ in both cases). The values of z in Session 3 and in Sessions 4 and 5 combined, while in the direction predicted by theory, only tended toward significance at the .05 level ($z = 1.79$ in Session 3 and $z = 1.92$ in Sessions 4 and 5 combined).

Despite the failure to attain significance for the later sessions, the general pattern of these results supports the proposition that early in acquisition, the probability of the avoidance response is directly related to the suppressing capacities of the tone.

The bottom section of Fig. 1 shows the course of emotionality in the posttone period on avoidance trials and on nonavoidance trials. A third assertion of dual process theory is that reinforcement for successful avoidance consists of a reduction in conditioned emotionality. It can be seen that, during the first few sessions, even when avoidance responses occurred, positively reinforced behavior tended to be suppressed during the posttone period. This finding raises the question of whether successful avoidance responses were actually accompanied by a reduction in emotionality and if so, at what point did this effect begin? Sign tests were also used to assess this question. However, in those tests, the paired items were the two suppression ratios (tone and posttone) from a given animal on a given trial where shock was avoided. The tests revealed that in Session 1, suppression during the posttone period was not reliably different from suppression during the tone; $z = 1.85$ ($P > .05$). From Session 2 on, however, with the occurrence of an avoidance response, the level of relative suppression underwent a statistically significant decrease. Each of the values of z, for Sessions 2 through 12 was greater than 3.74 ($P < .01$ for a two-tailed test, in each case). Thus, if reinforcement consists of a reduction in emotionality, these data suggest that the avoidance response was reinforced only from the second session on.

Figure 2 shows the percentage of avoidance response per block of five

trials and serves to illustrate the changes in performance that occurred within sessions. The within-session changes in emotionality are not shown, because when based on samples of only five trials, the random fluctuations of the several indices were of such magnitude as to obscure any underlying trends.

As seen in Fig. 2, during the final session of tone adaptation there were no systematic changes in the tendency to press the bar during the tone. With the introduction of shock at the end of tone, however, a performance emerged in which, during the initial five sessions, the tendency to avoid increased markedly within each session and decreased (to a lesser extent) in the 48-hr. interval between sessions. After Session 5, the performance had reached a stage in which very few shocks were received and such changes as occurred within and between sessions were small and unreliable.

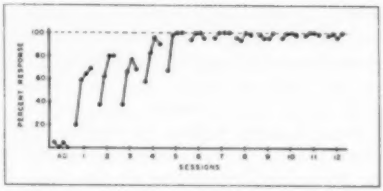


FIG. 2. Percentages of avoidance responses per block of five trials throughout the course of acquisition. (AD refers to the final session of tone adaptation.)

DISCUSSION

Interpretation of these results must recognize that the suppression ratio is, at best, an index of relative emotionality. It reflects the magnitude of a *change* in emotionality, but it provides no information about the absolute level of emotionality just prior to the change (i.e., the emotionality during the pretone period). Although none of the experimental measures in the present study were geared to provide an accurate assessment of the absolute level of emotionality, the data on the rate of plate presses during the pretone period is relevant to this question. During the final session of tone adaptation, the median pretone rate was 46 responses per min. During Session 1, when shocks occurred frequently, the median rate fell to 25 responses per min. However, as sessions progressed (and shock frequency decreased) the median pretone rate increased with each session, until by Session 7, it had reached 44 responses per min. Thereafter, the median pretone rate remained relatively constant, never falling below 40 responses per min., nor rising above 50 responses per min. From these data, it may be hypothesized that with the introduction of shock, the absolute level of emotional reactivity increased greatly, but that as sessions progressed, it slowly declined. Apparently, the obtained suppression ratios (represented in Fig. 1) reflect changes in emotionality over and above a moving baseline of generalized emotional reactivity. In this respect, it should be noted that a number of investigators have shown that some form of generalized emotional reactivity is an important variable in conditioning (Spence, 1958; Spence, Farber, & Taylor, 1954). Moreover, the results of a previous study by Hoffman, Fleshler, and Chorney (1961) cast additional light on this issue. In that study, it was found that even after extensive training, certain rats would fail to avoid on the early trials of each session, but would achieve a high level of performance by the end of the session. The results of that study indicated that the occurrence of shock was the critical factor in this warm-up-like phenomenon and for this reason suggested that warm-up in avoidance reflects a motivational process. Apparently, as shocks occur, their emotional aftereffects persist and summate to produce a state of emotional reactivity which facilitates avoidance.

In the present study, warm-up was exhibited during acquisition (Fig. 2) and in the absence of evidence to the contrary, it is reasonable to assume that this

feature of the performance also reflects the action of the lingering emotional aftereffects of aversive stimulation. Although existing dual process theories have not formally treated this particular process, Spence (1956) in dealing with classical conditioning, suggests that ". . . the drive level operating at the time of the conditioned anticipatory response is a function of the residual effects of the internal response (r_e) to the noxious stimulus of the preceding trials. That is, such emotional responses are assumed to have a relatively persisting effect that extends well beyond the range of temporal intervals usually employed in conditioning experiments . . ." (p. 186). It is clear that a motivational process, such as Spence describes, can be readily incorporated within existing dual process theory and hence that the occurrence of warm-up is consistent with a dual process interpretation of avoidance.

Summary

In order to examine the interplay between instrumental and emotional behavior during the acquisition of a discriminated avoidance response, rats were trained to press a bar to avoid shock while they were concurrently engaged in pressing a plate for food. The course of emotionality was tracked by assessing the several levels of suppression of ongoing plate presses during each of the various phases of the acquisition process. The results revealed a complex relationship between the level of performance on avoidance and the several concurrent indices of emotionality. In general, the results support the dual process hypothesis that conditioned emotionality controlled by the warning signal provides motivation for the avoidance response, while a decline in emotionality (with the offset of the signal) reinforces the response. The results also suggest that the lingering motivational aftereffects of aversive stimulation play an important role in the early phases of acquisition.

References

Black, A. H. The extinction of avoidance responses under curare. Unpublished doctoral dissertation, Harvard University, 1956.

Brady, J. V., & Hunt, H. An experimental approach to the analysis of emotional behavior. *J. Psychol.*, 1955, **40**, 313–325.

Estes, W. K., & Skinner, B. F. Some quantitative properties of anxiety. *J. exp. Psychol.*, 1941, **29**, 390–400.

Hoffman, H. S., Fleshler, M., & Chorney, H. Discriminated bar-press avoidance. *J. exp. Anal. Behav.*, 1961, **4**, 309–316.

Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.

Miller, N. E. Learnable drives and rewards. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951.

Mowrer, O. H. *Learning theory and personality dynamics.* New York: Ronald, 1950.

Solomon, R. L., & Brush, E. S. Experimentally derived conceptions of anxiety and aversion. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1956.* Lincoln: Univer. Nebraska Press, 1956. Pp. 212–305.

Solomon, R. L., & Wynne, L. C. Traumatic avoidance learning: Acquisition in normal dogs. *Psychol. Monogr.*, 1953, **67**(19, Whole No. 354).

Spence, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.

Spence, K. W. A theory of emotionally based drive (*D*) and its relation to performance in simple learning situations. *Amer. Psychologist*, 1958, **13**, 131–141.

Spence, K. W., Farber, I. E., & Taylor, J. The relation of electric shock and anxiety to level of performance in eyelid conditioning. *J. exp. Psychol.*, 1954, **48**, 404–408.

(Received August 14, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 295–299

ON THE RELATIONS AMONG SOME FACTORS THAT CONTRIBUTE TO ESTIMATES OF VERTICALITY

C. R. CURRAN AND H. L. LANE

University of Michigan

Recent research has identified some factors that effect nonveridical perception of verticality (Asch & Witkin, 1948a, 1948b; Wapner, Werner, & Chandler, 1951; Wapner, Werner, & Morant, 1951; Werner, Wapner, & Chandler, 1951; Witkin, 1949, 1950, 1952; Witkin & Asch, 1948a, 1948b). Through the experimental manipulation of visual context, body tilt and body support, among other variables, visual, proprioceptive, labyrinthine, and tactile cues have all been implicated as contributors to erroneous estimates of the upright. Further progress in the analysis of perception of the upright would seem to require at least the following three steps:

Isolation of the sensory events involved.—This could be achieved through experimental control, requiring manipulation of variables one at a time, or through partialing out the effects of these variables, requiring multidimensional experimental design, or by both techniques. Earlier studies have often confounded the effects of sensory events in several modalities when investigating the relation between a relatively complex experimental operation and perception of the upright. For example, the procedure of tilting *S* has been the most widely used experimental operation in research on perception of the upright, yet it may produce concurrent changes in visual, proprioceptive, labyrinthine, and tactile stimulation. Werner, Wapner, and Chandler (1951) have interpreted the effect of body tilt on estimates of the upright as evidence for their sensory-tonic field theory of perception, which suggests that the degree of muscular involvement plays an important role in determining these judgments. This inference must remain tentative, however, until the cluster of sensory changes effected by body tilt is experimentally analyzed. A similar case may be made for analyzing the complex effects of changes in the visual field. Witkin and Asch (1948b) have established that "the effect of the visual field upon the perceived upright tends to be stronger and more consistent the more richly articulated the field" (p. 782). However, richness of articulation is undoubtedly a multidimensional affair.

Use of independent variables measured on ratio scales.—Recent studies of perception of the upright have typically employed dependent variables measured on ratio scales (e.g., the angle between a rod called vertical by *S* and the true vertical) and one or more experimental treatments defined by nominal or ordinal scales (e.g., "tilted standing" and "tilted sitting"). Clusters of variables, such as support, lend themselves to nominal scaling; fractionation of these clusters should point to stimulus variables defined in terms of physical dimensions and thus measured (typically) on ratio scales. When *both* the dependent and independent variables are measured on ratio scales the predictive power of the findings is greatly enhanced, because the ratio scale contains the interval scale within itself (as well as the ordinal and nominal scales) (Stevens, 1960).

Quantitative analysis of the effects of variables and their interactions in multidimensional experimental design.—If a set of variables affecting perception of the upright has been isolated, if each can be measured on a ratio or, at least, interval scale, and if several levels or values of each variable are incorporated in an experimental design, then it is

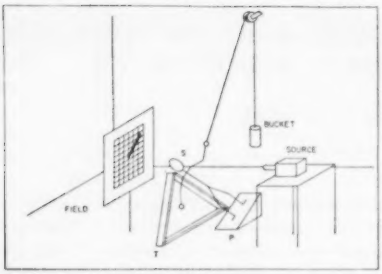


FIG. 1. Arrangement of apparatus for measurement of nonveridical perception of the upright.

possible to quantify their relative effects and their interaction effects in determining perception of the upright and to obtain the predictive power we desire.

The research to be reported represents a modest attempt to assess the effects of incorporating these three methodological improvements in an investigation of perception of the upright. The study describes the relation between changes in visual, proprioceptive, and labyrinthine stimulation and the degree of tilt of a rod, when it is reported to be vertical by the observer.

METHOD

The major experimental variables were: the illumination of the visual field (10^{-2}, 10^{-3}, or 10^{-4} ft-c), the degree of body tilt (10° or 30°), and the counterbalancing weight (equilibrium, [W], W plus 6 lb. and W plus 12 lb.). Parametric variables were: direction of tilt (left or right) and starting position of the rod (left or right).

Twelve male, naive undergraduates were *S*s in sessions lasting from 60 to 90 min. All *S*s had normal, uncorrected vision. Their heights ranged from 5 ft. 7 in. to 6 ft. 2 in., their weights from 131 to 179 lb. Earlier research has reported wide individual differences in judgments of the upright under distorting conditions (Witkin, 1949) but considerable consistency of judgment within individuals. A four-way experimental design was therefore employed, comprised of the three major variables (illumination, tilt, and counterbalancing weight) and a "blocking variable" (*S*s). Direction of tilt and starting position of the rod were confounded with *S*s, so that the four levels of the blocking variable were: right tilt, right rod ($N = 3$), left tilt, right rod ($N = 3$), etc. Each *S* gave 3 judgments under each of 18 combinations of the levels of illumination, tilt, and weight, presented in counterbalanced order.

Figure 1 is a schematic representation of the apparatus. A triangle (T) with base angles 60° and 80°, was constructed from 2 × 4 in. boards and served as a tilt reference for *E*. The *E* aligned the median plane of *S* with the appropriate arm of the tilt reference by visual inspection before each trial. With *S* standing erect, the light source, 4 ft. behind him, and the center of the stimulus field, 4 ft. in front, were at eye level and in approximately the same vertical plane. The *S* stood on a pedestal (P) that was held firmly in place on the floor. The surface of the pedestal slanted upward at an angle that was set equal to the angle of tilt, so that *S* was perpendicular to the surface when aligned with the tilt reference. When *S* was tilted, he grasped the supporting rope in the hand contralateral to the direction of tilt. The rope passed over a pulley suspended from the ceiling (18 ft. high) and terminated in a bucket of cement, selected by *E*.

This cement weight was selected in the following way. At 10° tilt, a weight, W_{10}, was determined that was within .5 lb. of 3% of *S*'s weight. At 30° tilt, $W_{30} = 25\%$ of *S*'s weight, ±.5 lb. Under either tilt condition, the levels of the counterbalancing variable were then W, W + 6 lb., and W + 12 lb. The value of W under each condition of tilt approximated the weight that would just balance *S*. The moments of force around the pedestal are approximated by: $M(h_c \sin T) = Wh_a$ where M = weight of *S*, h_c = distance from the fulcrum to *S*'s center of gravity, T = angle of tilt from the vertical, W = counterbalancing weight, and h_a = distance from the fulcrum to *S*'s arm.

Under all experimental conditions, *S* viewed the stimulus field monocularly, through a reduction tube (3 in. long, 2 in. in diameter) attached to a pair of goggles. In this manner, the visual field of the eye contralateral to the direction of tilt was restricted to the stimulus field, while vision in the homolateral eye was blocked.

The field had an illuminance of approximately 10^{-2} ft-c with unfiltered illumination from a source whose color temperature was approximately 1800°K (34 v. applied to a

500-w. lantern slide projector). Wratten neutral density filters were inserted to produce illuminances of 10^{-3} and 10^{-4} ft-c giving the three levels of the illumination variable. The stimulus field consisted of a piece of white poster board, 60 × 40 in., with a luminous reflectance of approximately 0.83. A piece of linear graph paper, 8 × 8 in., with blue lines heavy-ruled at 1-in. intervals and light-ruled at .1-in. intervals, was mounted on the front of the poster board and aligned with the true vertical by means of a plumb line. (The luminous reflectance of the graph paper was approximately 0.77; at an illuminance of 10^{-4} ft-c no *S* reported seeing the rulings while at 10^{-2} ft-c the paper and its rulings were visible to all *S*s.) Behind the poster board, a small motor rotated a shaft that punctured the board at its center and protruded ½ in. A brass rod (reflectance 0.40) 12 in. long, ¼ in. diameter, was mounted at right angles to this shaft in front of the stimulus field for observation by *S*, and a 26-in. rod was mounted in parallel behind the poster board for observation by *E*. Twenty-four inches above the pivotal point of the rod at the center of the field, *E* could read the point of intersection of the rod with a horizontal line, ticked at ¼-in. intervals; he could therefore read the position of the rod to an accuracy of about 0.5°. An ac motor with reduction gears swept the rod from its starting position, 50° ± 10° to the left or to the right of the vertical, toward the upright at a constant rate of 1.5° per sec. When *S*, who was observing the position of the rod under a given condition of tilt, weight, and illumination, judged it to be vertical, he said "stop" aloud, at which time *E* pressed a normally closed switch to stop the motor. He then read the rotary position of the rod, using a pencil light to read the horizontal scale. Following a trial, *S* was instructed to close both eyes while *E* changed tilt, weight, illumination, and rod position, in accordance with the protocol.

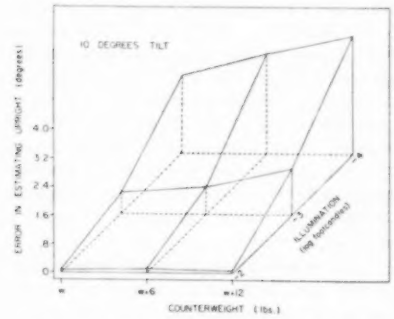


FIG. 2. The effect of counterbalancing weight, illumination, and tilt (10°) on perception of the upright.

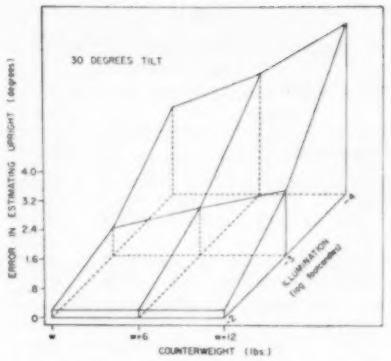


FIG. 3. The effect of counterbalancing weight, illumination, and tilt (30°) on perception of the upright.

RESULTS AND DISCUSSION

Figures 2 and 3 summarize the findings of this experiment by depicting the effect of the weight and illumination variables on estimates of the upright at each of the two levels of the tilt variable. Examination of these figures reveals the relative contribution of these three variables as well as their interaction effects. Clearly, the major variable is the level of illumination: the error solid grows rapidly in volume as the level of illumination is decreased at 10° tilt and even more rapidly at 30° tilt; this comparison reveals an Illumination × Tilt interaction. At relatively high levels of illumination the counterbalancing weight has little or no effect at 10° tilt and only a slight effect at 30° tilt. As illumination is decreased, however, the counterbalancing weight plays an increasing role in determining nonveridical perception of the upright. Comparison of Fig. 2 and 3 reveals a Weight × Tilt as well as a Weight

TABLE 1

ANALYSIS OF VARIANCE OF ESTIMATES OF THE UPRIGHT

Source	*df*	*F*
Illumination (I)	2	525.26**
Tilt (T)	1	72.15**
Counterweight (C)	2	56.74**
I × C	4	5.23**
C × T	2	4.52*
I × T	2	3.20*
I × C × T	4	4.33**
Within cells (*MS*)	576	(0.85)

* $P < .05$.
** $P < .01$.

× Illumination interaction, and a Weight × Tilt × Illumination interaction. The solid obtained at 30° tilt is appreciably larger in all cells than that obtained at 10° tilt, showing the net effect of the tilt variable. As just indicated, there are also obvious Tilt × Weight and Tilt × Illumination interactions.

Table 1 presents an analysis of variance of the estimates of verticality. As anticipated, the variance attributable to replications within *Ss* is relatively small. All three main effects and their first- and second-order interactions are significant at the .05 level or beyond.

The present findings support Witkin's (1949) analysis of the relative contribution of visual as opposed to somesthetic cues in the determination of perception of the upright. These findings show, furthermore, that relatively few visual cues to the vertical can yield extremely accurate estimates of the vertical, even under marginal visibility, with the body tilted and delicately poised. It appears that the visual field need not be "richly articulated" to permit accurate estimates of the upright under distorting conditions. However, removal of these few visual cues by approximately halving the brightness of the field yielded almost a hundredfold increase in error in perceiving the upright.

When visual cues were minimized, at the lowest illuminance, and proprioceptive and cutaneous cues were minimized in the condition of equal moments around the pedestal, the degree of tilt was observed to have a considerable effect on the magnitude of the error in perceiving the upright. Inference from these findings suggests that, in the absence of visual and somesthetic cues to the upright, other cues to the static position of the body are available; perhaps the utricular otoliths, which are thought to play a role in static positional adjustments of the body (Geldard, 1953, p. 262) are the source of this stimulation.

When the condition of equilibrium is displaced through the addition of counterbalancing weights, greater effort is required on the part of *S* to maintain his balance and his alignment with the tilt reference. As the counterbalancing weight is increased, the magnitude of the error in judging the vertical is increased. This finding may be related to the qualitative prediction of Werner, Wapner, and Chandler (1951): "Within the framework of the sensory-tonic field theory of perception, the degree of muscular involvement is expected to be an important variable" (p. 346). The results of the present study also confirm the observation of these authors that the apparent vertical is shifted to the side opposite the direction of the body tilt. It will be remembered that the direction of tilt was confounded with the starting position of the rod and with *Ss* and incorporated into the experimental design as a four-level blocking variable. A posteriori comparison of the weighted means of the right rod, right tilt and left rod, left tilt *Ss* with the weighted means of the right rod, left tilt and left rod, right tilt *Ss* showed that homolateral tilt and rod starting positions produce a greater magnitude of error than contralateral positions of these variables (Scheffé's method; $\alpha = .05$).

It is noteworthy that all interaction effects are large and significant. Although these interaction effects have not been demonstrated previously in the analysis of perception of the upright, this seems a plausible way for a perceptual

process to be controlled by the relevant variables. As indicated earlier, this type of analysis is facilitated through the use of multidimensional experiments with unidimensional variables sampled at several levels and measured on ratio scales.

Summary

Several variables that have been shown to influence the perception of the upright were incorporated in a multidimensional design to permit analysis of their several effects and interactions. Minimal visual cues had a dramatic effect in reducing nonveridical perception of the vertical. Distortion of body tilt and balance produced effects of lesser magnitude. All the first- and second-order interactions of these variables had large and significant effects on perception of the upright.

References

Asch, S. E., & Witkin, H. A. Studies in space orientation: I. Perception of the upright with displaced visual field. *J. exp. Psychol.*, 1948, **38**, 325–337. (a)

Asch, S. E., & Witkin, H. A. Studies in space orientation: II. Perception of the upright with displaced visual fields and body tilted. *J. exp. Psychol.*, 1948, **38**, 455–477. (b)

Geldard, F. A. *The human senses.* New York: Wiley, 1953.

Stevens, S. S. Ratio scales, partition scales, and confusion scales. In H. Gulliksen and S. Messick, (Eds.), *Psychological scaling: Theory and application.* New York: Wiley, 1960.

Wapner, S., Werner, H., & Chandler, K. Experiments on the sensory-tonic field theory of perception: I. Effect of extraneous stimulation on the visual perception of verticality. *J. exp. Psychol.*, 1951, **42**, 341–345.

Wapner, S., Werner, H., & Morant, R. B. Experiments on the sensory-tonic field theory of perception: III. Effect of body rotation on the visual perception of verticality. *J. exp. Psychol.*, 1951, **42**, 351–357.

Werner, H., Wapner, S., & Chandler, K. Experiments on the sensory-tonic field theory of perception: II. Effect of supported and unsupported tilt of the body on visual perception of verticality. *J. exp. Psychol.*, 1951, **42**, 346–350.

Witkin, H. A. Perception of body position and of the position of the visual field. *Psychol. Monogr.*, 1949, **63**(7, Whole No. 302).

Witkin, H. A. Perception of the upright when the direction of the force acting on the body is changed. *J. exp. Psychol.*, 1950, **40**, 93–106.

Witkin, H. A. Further studies of perception of the upright when the direction of the force acting on the body is changed. *J. exp. Psychol.*, 1952, **43**, 9–20.

Witkin, H. A., & Asch, S. E. Studies in space orientation: III. Perception of the upright in the absence of a visual field. *J. exp. Psychol.*, 1948, **38**, 603–614. (a)

Witkin, H. A., & Asch, S. E. Studies in space orientation: IV. Further experiments on the perception of the upright with displaced visual fields. *J. exp. Psychol.*, 1948, **38**, 762–782. (b)

(Received August 15, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 300-310

THE PERSPECTIVE ILLUSION: PERCEIVED SIZE AND DISTANCE IN FIELDS VARYING IN SUGGESTED DEPTH, IN CHILDREN AND ADULTS[1]

JOACHIM F. WOHLWILL

Clark University

Since the Renaissance it has been generally known that perspective drawings can convey a strong impression of depth in two-dimensional stimulus fields; curiously, however, this phenomenon has received little systematic attention on the part of psychologists. To be sure, Gibson, in his treatment of space perception, and of pictorial perception generally, has repeatedly pointed to the important role of perspective (Gibson, 1950, 1954, 1960); yet, apart from a largely exploratory study by Smith, Smith, and Hubbard (1958), the actual effects of perspective on perception have not been experimentally investigated.

The present study is concerned with one aspect of this problem, viz. the extent to which stimulus fields constructed according to the principles of perspective geometry will affect judgments of size and length within such a field. Thus, one might expect the height of an object located at the bottom (i.e., apparent front) of a perspective drawing to be perceived as smaller relative to a similar object located at the top (i.e., apparent rear) of the drawing. In fact, just such an essentially illusory effect is dramatically illustrated by Gibson (1950, p. 182), by means of a perspective drawing of a corridor in which several barrel-shaped objects are depicted: the rearmost barrel appears strikingly expanded in size in comparison with the front one. More systematic investigation of these effects, and especially of their variation as a function of the characteristics of the stimulus fields responsible for the suggestion of depth should provide a clearer picture of this illusion, its magnitude and its determinants. Such a study should furthermore be of direct relevance to Gibson's (1950) general theory of space perception, which emphasizes the information to depth contained in the gradients of texture-density, etc., present in any two-dimensional projection of a three-dimensional stimulus field, be the projection retinal, photographic, or in the form of a perspective drawing.

The aim of this study is thus to investigate experimentally the effects of perspective drawings on the perception of relative linear extent in the plane of the drawing, with reference both to the perceived size of objects in this plane and the perceived distance between points in the plane. The principal variable manipulated in

[1] This investigation was supported by a grant from the National Science Foundation, G-16031. The author wishes to thank Lawrence Houle, Principal of the Columbus Park School in Worcester, Massachusetts, and the teachers of this school for their cooperation in supplying subjects and facilities for the experimental work at the grade levels. The helpful comments of Keith Smith and Benjamin White of the MIT Lincoln Laboratories, regarding the informational analysis presented at the end of this paper, are also gratefully acknowledged. Finally, Peter Schiller provided valuable assistance in the design and construction of the apparatus.

the study is the nature of the "information to depth" contained in the field: first, the greater the amount of this information, as expressed in the *density* of the texture of the field, the greater should be the distorting effect of perspective; second, the introduction of redundancy into the field, expressed in terms of the patterning of the elements of texture, so as to enhance linear perspective, should likewise increase the effect. The effect should be maximal, finally, in a field portraying directly the geometrical relationships involved in a perspective transformation.

An additional variable of considerable interest in this domain of perception is that of the age of *S*s. There is considerable evidence that spatial relationships generally exert relatively little influence on the perception of young children (cf. Wohlwill, 1960), so that one might postulate that the distorting effect of perspective is absent, or at least rather small in magnitude in early childhood, and will increase with age. Indeed, Glasser

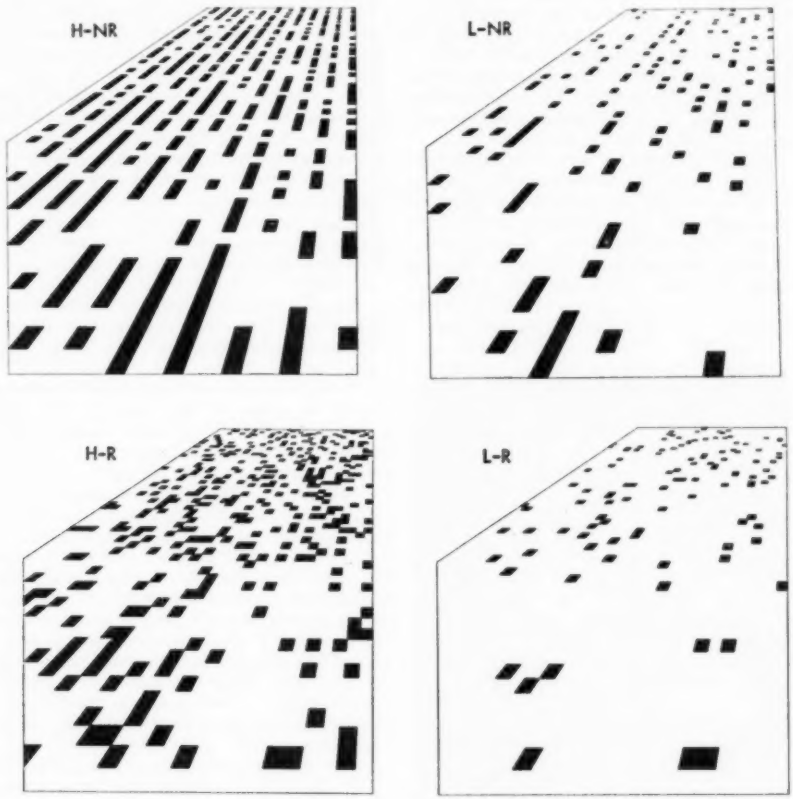


FIG. 1. Perspective drawings representing four of the stimulus fields utilized in the study, with texture-density (left vs. right figures) and randomness (top vs. bottom figures) as variables.

(1944) has claimed that young children rarely perceive a perspective illusion in a figure very similar to that of Gibson referred to above, and accordingly attributes the effect to the role of experience; however, no data are adduced in support of this statement, nor is the age of *S*s to which it is intended to apply further specified.

METHOD

Stimuli.—Six different stimulus fields, drawn on sheets of drafting paper 23 × 29 in. in size, were employed in this study. The shape of all of these fields was uniform, consisting of a trapezoid superimposed on a rectangle; the bases of the trapezoid were 7⅛ in. and 12 in. and its height 7 in., while the dimensions of the rectangle were 12 × 18 in. The intersection of the two sides of the trapezoid at a point 12 in. above its lower base defined the vanishing point used for constructing the perspective drawings.

These fields were filled with different perspective drawings, all of which were based on a grid of 36 columns fanning out from the vanishing point and 62 rows spaced so as to produce the foreshortening of distance required by the laws of perspective. The grid itself, with lines of uniform thickness drawn in India ink, made up the first panel (illustrated in Fig. 2, below). Four of the remaining panels, shown in Fig. 1, were constructed by filling in selected cells from this grid in India ink, according to the following plan: (*a*) High density, nonrandom (Fig. 1, top left): Every third column was selected from the grid; within each the cells to be filled in were determined by a table of random numbers, so as to yield an average proportion of ⅔ of the cells filled in per column, or ⅓ of the total number of cells in the field. (*b*) Low density, nonrandom (Fig. 1, top right): From among those cells selected under *a*, a subset was chosen by means of a table of random numbers, consisting of $\frac{3}{10}$ of the cells of *a*, or a density of .18 per column, i.e., .06 for the total field. (*c*) High density, random (Fig. 1, bottom left): From every column of the grid ⅓ of the cells were randomly selected to be filled in. (*d*) Low density, random (Fig. 1, bottom right): From among the cells selected for *c*, a subset consisting of $\frac{3}{10}$ of these cells, i.e., .06 of the cells of the total field, was chosen.

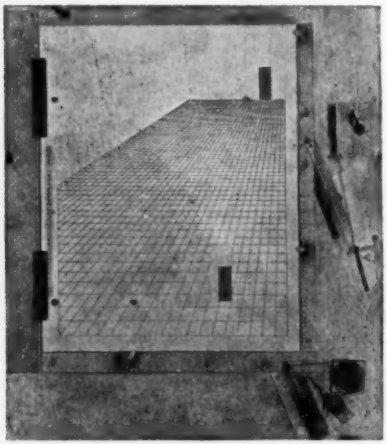


FIG. 2. Apparatus, with grid panel displayed.

It should be noted that the proportion of cells from the total grid that were filled in was the same for both high-density panels (.20), as well as for both low-density panels (.06). Further, it will be seen that the essential difference between the random and the nonrandom panels lies in the greater sense of linear perspective afforded by restricting the cells to a limited set of columns, in the case of the latter.

Finally, the sixth panel was a control field, which was entirely blank, except for the border of the field, drawn in in India ink as in the other panels.

These six panels will henceforth be referred to as: G (Grid); H-NR (high-density, nonrandom); L-NR (low-density, nonrandom); H-R (high-density, random); L-R (low-density, random); and C (control).

Apparatus.—The apparatus used for displaying the stimulus panels and for manipulating the size and distance variables is shown in Fig. 2. (Figure 2 also shows Panel G, as it was exposed in the apparatus.) The apparatus consisted essentially of a rectangular piece of plywood, 29.5 × 35 in., covered by a sheet of Plexiglas fitted into a frame which was attached by means of hinges to the side of the plywood base. The Plexiglas cover could thus be opened for the insertion and removal of the individual panels exposed behind it.

For the *distance* judgments, two points were chosen in the stimulus field of the panels, defining the distance to be bisected by *S*. The points were located on an imaginary line passing through the vanishing point of the

perspective drawings, one point towards the top edge, the other near the bottom. Screws were driven through two corresponding spots on the Plexiglas cover, directly superimposed on the two points in the stimulus field as it was exposed underneath the cover.

An endless loop of fine nylon thread ran around a trapezoidal path marked by these two screws and two others near the left side of the Plexiglas cover and at the same heights as the former pair. This loop could be moved in either direction by pulling on a knot located along the vertical side, this movement causing a little red plasticine ball to travel up and down between the first two screws. A small indicator attached to the left vertical segment of the line, which moved along a ruler glued to the Plexiglas, enabled *E* to read off the height of the ball corresponding to the midpoint of the distance as perceived by *S*.

For the *size* judgments, two rectangular cutouts were made in each panel, through which the stimulus objects—two light-blue rectangular sheets of metal—appeared. One of the cutouts was at the top of the panel, the other towards the bottom, their lower-left vertices being located along a line through the vanishing point, well to the right of the line involved in the distance judgments (cf. Fig. 2).

The stimulus objects were attached to two wooden slides which could be moved up or down, along a line which was an extension of the diagonal of the rectangles, i.e., in an oblique direction relative to the field of the panel; movement of these slides thus caused the portion of the blue rectangle exposed through the cutout to vary in size, but with a constant ratio of height to width. Each slide was raised and lowered by turning a crank, to which it was connected by a set of pulleys.

This arrangement permitted continuous variation of both the top and bottom stimulus objects, so that either could be used as a variable and the other as a standard. The size of the standard rectangle, in terms of the length of the diagonal, was 7.5 cm.; the diagonal of the variable rectangle varied from 0 to 12 cm. In order to prevent *S*s from utilizing the amount of white in the cutouts around the rectangles as a cue, the two cutouts were made unequal in size, the bottom one being 4.3 × 11.3 cm. (a size corresponding to the maximum size of the variable), whereas the top one measured 5.3 × 12.3 cm.

The whole apparatus was displayed in a vertical position to the *S*, by sliding it into a metal frame mounted vertically on a wooden base.

Procedure.—The adult *S*s were tested in a room illuminated only by overhead fluorescent lighting. They were seated on a stool at a distance of 3 m. from the apparatus, so that their eyelevel was approximately even with the center of the stimulus panel. The *S* was told that the experiment concerned his ability to make judgments of size and distance relationships. Specifically, for the *distance* judgments he was told that *E* would make the little red dot travel upward along the line between the two screws, and instructed to say "stop" when he thought the dot was exactly midway between the screws. This wording of the instructions was intended to foster a set for objective, rather than phenomenal judgments; however, *E* discouraged *S* from attempting to "figure out" where the midpoint was intellectually. The *S* was further informed that he would have an opportunity to correct any setting if he was dissatisfied with it.

For the *size* judgments the instructions were similar, *S* being asked to say "stop" when the top (or bottom) rectangle was just equal in size to the bottom (or top) one.

In his manipulation of the red dot and of the variable rectangle, *E* always faced away from *S*, so as to avoid giving him any involuntary facial cues that might influence his judgment. Movement of the stimuli was carried out at a fairly even rate, although for the variable size stimulus perfect smoothness was not realizable with the apparatus as constructed.

Special procedures were used with the group of children from Grade 1, in order to ensure that they properly understood the instructions. These procedures consisted of a series of pretest judgments, involving, (*a*) marking the middle of two lines drawn on a 8½ × 11 in. sheet of paper, one line horizontal, the other oblique, (*b*) deciding when a bead which *E* moved along a string over the surface of a table, first parallel to the edge, then obliquely, was at the middle of the string, and (*c*) making a preliminary distance setting on the apparatus, without any panel underneath the Plexiglas cover. There were no instances in which any child gave an indication of failing to understand the instructions, either through questions directed to *E*, or through markedly deviant judgments on the pretest.

All of the children were routinely asked, "Is that it?" or "Is that where you want it?" after every judgment, and allowed to change it if they wished (as the adults had been also). This was done in an effort to minimize errors of anticipation. The children were further asked, after being presented with the

TABLE 1

MEAN DISTANCE SETTINGS (IN CM.)

Groups	Stimulus Panels						Combined	
	C	L-R	H-R	L-NR	H-NR	G	Mean	s^a
Grade 1	21.91	22.19	22.45	22.23	22.97	23.12	22.48	1.35
Grade 4	22.00	22.71	22.59	22.85	22.96	23.02	22.69	1.32
Grade 8	21.40	22.01	22.00	22.27	22.44	22.72	22.14	1.15
Adults	21.06	21.08	21.20	21.19	21.43	21.80	21.30	0.86
Combined Mean	21.59	22.00	22.06	22.13	22.45	22.67		
s^b	1.42	1.81	1.78	1.82	1.87	1.77		

Note.—Values tabled represent bisections of a 41-cm. vertical distance.
[a] *SD* of *Ss*' mean scores, based on between-*Ss* error terms for each age level.
[b] *SD* of scores for all *Ss*, based on residual error term calculated for each stimulus panel.

H-R and H-NR panels, whether they could think of anything that might look like what they were seeing.

Design.—For both distance and size judgments *S* made two judgments for each stimulus panel, one ascending, the other descending. Size and distance judgments were always made consecutively for any stimulus panel, before a new panel was exposed; half of the *Ss* always judged size first, the other half judged distance first. The six panels were presented in a Latin square design, in two different sequences, one being the reverse of the other. Regardless of the particular panels exposed, all *Ss* started with the ascending judgment on the first, third, and fifth panels and with the descending judgment on the second, fourth, and sixth panels. Finally, on the size judgments half of the *Ss* were tested with the top stimulus as the standard, and the other half with the bottom stimulus as the standard.

Subjects.—There were four groups of 24 *Ss* each, representing samples of children from Grades 1, 4, and 8 and college-age adults. The mean ages of these four groups in years and months were 7:1, 9:10, 14:0, and (approximately) 20:0, respectively. The school children all came from a lower-middle class grade school, and were thus not strictly comparable in IQ and related variables to the adults, who were college undergraduates (mostly freshmen and sophomores) enrolled in an introductory psychology course. All *Ss* reported they had normal eyesight, either uncorrected or corrected. (For the youngest children, the pretest with the apparatus given before the experiment proper allowed *E* to satisfy himself of the adequacy of *S*'s eyesight.)

RESULTS

Distance

The distance between the two screws which was to be bisected was 47.3 cm. The vertical component of this distance (on which the recorded distance data are based) was 41.0 cm. Thus settings of the red dot larger than 20.5, the objective midpoint, would indicate an influence of suggested depth. This influence is apparent in the judgments at all ages and for all stimulus conditions, including Panel C, as shown by the means shown in Table 1. (Additional comparisons between the H and L panels and between the NR and R panels are provided in Table 2.)

TABLE 2

MEAN DIFFERENCES IN DISTANCE SETTINGS (IN CM.) BETWEEN PANELS VARYING IN DENSITY AND IN RANDOMNESS

Groups	H vs. L	NR vs. R
Grade 1	.50	.28
Grade 4	.00	.26
Grade 8	.08	.35
Adults	.18	.17
Combined	.19	.26

Note.—Positive values indicate greater magnitude of illusion for first-listed panels.

As the analysis of variance of these data is rather lengthy and complex, suffice it to present it in abbreviated form. It falls into four parts. The first, involving all between-*S*s effects (age, order, and their interaction), disclosed the variance due to age to be significant at better than the .01 level ($F = 5.95$, $df = 3/88$). The second, involving within-*S*s effects summed over direction (stimuli, singly and in interaction with age and order), showed the variance due to the stimulus fields to be highly significant ($F = 14.92$, $df = 5/440$), but no significant interactions. The third part, involving within-*S*s effects summed over stimulus panels (direction, singly and in interaction with age and order) showed a highly significant effect due to ascending vs. descending direction ($F = 152.54$, $df = 1/88$), as well as an interaction between direction and age significant at between the .05 and .01 levels ($F = 3.05$, $df = 3/88$). The fourth part, finally, is comprised of within-*S*s effects due to simple and higher-order interactions involving both the stimulus and the direction variable; here the simple interaction was significant at the .01 level ($F = 3.09$, $df = 5/440$).

These results may now be summarized as follows:

Stimulus fields.—The observed differences between stimulus fields are in good agreement with those postulated in our introduction: the effect of perspective was greatest for the G panel and least for the C panel (cf. Table 1); it was also greater for high- than for low-density fields, and greater for nonrandom than for random fields (cf. Table 2). Duncan's multiple range test (cf. Edwards, 1960, pp. 136 ff.) indicates the following comparisons between means (for all groups combined) to be significant at the .05 level: G vs. all others except H-NR; H-NR vs. H-R, L-R, and C; L-NR vs. C; H-R vs. C; L-R vs. C. In addition, orthogonal comparisons (Edwards, 1960, pp. 140 ff.) between the two H vs. the two L fields, as well as between the two NR vs. the two R fields, both show differences significant at better than the .01 level.

Age.—The overall effect appears to *decrease* with age, except that Grade 4 *S*s showed slightly (but nonsignificantly) higher mean values than Grade 1 *S*s. An application of Duncan's multiple range test, however, shows that all three of the children's groups are significantly differentiated from the adults, but not from one another. Although there was a suggestion that the first graders were somewhat more influenced by the density variable than the other groups (cf. Table 2), the interaction between age and stimulus panels was not significant.

Direction.—A notable feature of the results was the finding of a very marked anticipation effect: for all age groups and stimuli combined, the mean ascending judgment was 21.44, while the mean descending judgment was 22.87. This effect itself interacted with age, being largest at the first- and fourth-grade levels, and considerably reduced at the eighth-grade and adult levels. There was also a significant interaction of this factor with the stimulus variable which was less consistent in nature.

Order.—The order-of-judgment variable (distance before vs. after size) failed to account for any significant portion of the variance, either singly or in interaction.

Size

The size of the standard rectangle (measured in terms of the length of the diagonal) was 7.5 cm. Thus settings of the variable smaller than 7.5,

TABLE 3

MEAN SIZE MATCHES, IN CM. (POE = 7.5)

Groups	Stimulus Panels						Combined	
	C	L-R	H-R	L-NR	H-NR	G	Mean	s^a
Grade 1	7.47	7.51	7.49	7.27	7.36	7.23	7.39	0.28
Grade 4	7.39	7.49	7.53	7.35	7.33	7.36	7.41	0.29
Grade 8	7.42	7.49	7.50	7.26	7.39	7.27	7.39	0.24
Adults	7.29	7.50	7.61	7.38	7.33	7.31	7.40	0.23
Combined	7.39	7.50	7.53	7.32	7.36	7.29		
s^b	0.32	0.34	0.39	0.35	0.37	0.41		

Note.—Values tabled represent height at top of field judged equal to a 7.5-cm. height at bottom (see text for details).
[a] See Footnote a, Table 1.
[b] See Footnote b, Table 1.

when the standard was at the bottom of the field, and larger than 7.5, when the standard was at the top of the field, would indicate an overestimation of the top stimulus, hence a perspective effect. In order to make the measures for the two positions of the standard comparable to each other, the settings obtained with the standard at the top were translated into scores which represented the size of the top rectangle perceptually equivalent to a 7.5-cm. rectangle at the bottom. This was accomplished by means of the formula $s'/7.5 = 7.5/s$, where s and s′ represent, respectively, the match made to the 7.5 standard at the bottom, and the transformed score.

TABLE 4

MEAN DIFFERENCES IN SIZE MATCHES (IN CM.) BETWEEN PANELS VARYING IN DENSITY AND IN RANDOMNESS

Groups	H vs. L	NR vs. R
Grade 1	.04	−.18
Grade 4	.01	−.17
Grade 8	.07	−.17
Adults	.03	−.18
Combined	.04	−.18

Note.—Negative values indicate greater magnitude of illusion for first-listed panels.

In order to simplify somewhat the analysis of variance (complicated even further, beyond the already rather unwieldy one dealt with for the distance judgments, due to the addition of the variable of the position of standard) each *S*'s settings for the ascending and descending conditions for each stimulus panel were averaged.

The means for each stimulus field at each age level are shown in Table 3, while Table 4 provides a comparison of the H and L panels and the NR and R panels. The analysis of variance of the data falls into two parts. The first, between-*S*s portion (comprising the variables of age, order and position of standard, singly and in interaction) discloses no significant source of variance. The second, within-*S*s portion (comprising the stimulus variable, singly and in interaction with the others), shows a significant effect due to the stimulus fields ($F = 7.66$, $df = 5/400$, $P < .01$), but no significant interactions.

Summarizing and at the same time elucidating these results, we find the following:

Stimulus fields.—While this variable had a significant effect on the size judgments, the results were much less consistent, and less closely in agree-

ment with expectations, than was the case for the distance judgments (cf. Table 3). Duncan's range test shows that the means for both H-R and L-R conditions were significantly higher (at the .01 level) than those for the H-NR, L-NR, and G conditions. This was what had been anticipated, since for these judgments, the lower the score, the larger the apparent size of the top stimulus, and hence the greater the magnitude of the illusion. The results for the C (Control) condition are, however, decidedly out of line, since instead of yielding the highest mean, this condition emerges as intermediate between the two random and the two nonrandom stimulus fields; in fact, the difference between H-R and C is significant at the .01 level (in the wrong direction)! Furthermore, while the two nonrandom-panel means clearly differ from the two random-panel means as expected, the two high-density means are both higher (though not significantly) than the two low-density means, whereas the opposite was anticipated (cf. Table 4).

Age.—There appeared to be no consistent differences between the age groups, nor did this variable interact with the stimulus variable.

Direction.—Although this variable did not enter into the analysis of variance, inspection of the data shows again a very marked anticipation effect. This effect also tended to decrease with age, but the main difference appeared to be between the adults on the one hand and the three groups of children on the other.

Position of standard.—The mean for all judgments made with the variable at the top was 7.44, as against a mean of 7.36 for the transformed settings made with the standard at the top. While this difference was not statistically significant ($F = 2.21$, $df = 1/80$), it is in the direction of the "error of the standard," involving an overestimation of the standard stimulus per se, which has been encountered previously in the literature (Gardner & Long, 1960; Piaget & Lambercier, 1943). The interaction of this effect with age likewise was nonsignificant; it might be noted nevertheless that the effect appeared to be most marked at the fourth grade, while the eighth graders and adults failed to exhibit any trace of it.

Order.—The order-of-judgment variable again failed to affect the judgments significantly.

Verbalizations Given to the H-R and H-NR Stimulus Fields

Of the 72 school children asked for an interpretation of the H-R and H-NR fields, 43 responded to the H-R and 48 to the H-NR panels. Of these 91 responses, 78 clearly referred to a scene seen in depth, the most common response being a floor, or some variant thereof (e.g., a highway). Interestingly enough, 9 responses made reference to a vertical plane (skyscraper, building), of which 6 came from the eighth graders. Otherwise no notable age differences were found; even the failure-to-respond rate was essentially the same for all groups. Nor were there any very consistent differences between the two panels: the H-NR panel elicited substantially more depth responses on the part of the eighth graders, but the two younger groups gave slightly more depth responses to the H-R panel.

All in all, the results suggest that perspective drawings of this type are effective in conveying a sense of phenomenal depth even to the youngest of the *Ss* included in this study—a conclusion which is in line with the observed effect of these drawings on the perceptual judgments.

Discussion

The discussion of the results of this investigation will focus on four separate points:

1. With respect to the distance judgments, the effects of the perspective drawings conformed very neatly to our expectations, showing that information to depth contained in the field, as manipulated in this study, represents a major determinant of the perceptual distortions produced in such drawings. The size judgments, however, yielded much less conclusive results, and did not appear to be consistently related to the stimulus-information variable.

In attempting to account for these somewhat discrepant results, it may be helpful to examine the way in which *S* would in fact let the stimulus field affect him in making his judgment. It is apparent that for the distance judgments the background forms an essential part of *S*'s field of attention, since his task, to bisect the distance between the two screws, requires him to scan back and forth along a considerable portion of the stimulus field. In the case of the size judgments, on the other hand, it would be much easier for *S* to ignore the field separating the two stimulus objects to be compared. The incisions made into the field around the objects, and the location of the upper stimulus standing on the top border of the field may also have contributed to the perceptual isolation of the stimuli from the background fields, which would have mitigated their influence. Admittedly, it is difficult to account on this basis for the significant differences between some of the stimulus fields that were found, unless one assumes that, for whatever reason, the intrusion of the field into the size comparisons was especially slight for the two randomly textured panels.

2. The results with respect to the age variable can only be regarded as inconclusive. Certainly there was little suggestion of an increase in the effects of perspective with age, except for the very slight and nonsignificant increase from the first to the fourth grade. It seems, therefore, that if this illusion is a product of learning, whether in the sense of associative or assumptive processes or in the sense of increasing experience in responding to spatial relationships within a field, such learning must run its course fairly early in life. The significant *decrease* in the effects of perspective on the distance judgments for the adults, on the other hand, appear to reflect a more active attempt on the part of these *S*s to counteract this illusion, of which most of them were well aware. In this connection it should be noted that the difference between adults and children may have been at least in part a matter of intellectual level, rather than age, the adults clearly representing a more select group in this respect than the children.

3. An interesting finding was that for both distance and size judgments the control panel itself gave rise to a constant error in the direction of the perspective illusion. A variety of factors could have contributed to this result: the border of the field present in the control panel might by itself have conveyed a sense of depth; perseverative effects from other panels previously exposed might have led *S* to perceive the control panel in terms of depth; relative height in the field may be interpreted as depth, even in the absence of other cues, just as in other studies (Smith, 1958; Weinstein, 1957) this variable has been shown to represent an effective cue by itself, leading to a certain amount of constancy in size judgments made from photographs.

4. Finally, since one might consider this study as representing a beginning towards an informational approach to the study of space perception, a brief analysis of the possibilities as well as the limitations of such an approach in this area appear to be warranted.

Given the selection procedures employed in the construction of the stimulus fields, the specification of their formal informational content (independent of their role in suggesting depth) is relatively straight forward. As regards the density variable, the two random fields, for which p (the probability of a cell's being filled in) was, respectively, .06 and .20, contain an average of .34 and .72 bits per cell,

respectively. Similarly the two NR panels, for which p, for the restricted set of columns from which cells were selected, was .18 and .60, contain, respectively, .69 and .97 bits per cell for each filled column, or .23 and .32 bits per cell for the total field.

The regularity variable, on the other hand, can best be expressed in terms of the redundancy, relative to the corresponding R panels, introduced by the selection procedures. Thus, taking the R panels as a baseline, L-NR would show a redundancy of 32% (1−.23/.34), while that of H-NR would be 55% (1−.32/.72). It is interesting to note that for the distance judgments the effect of redundancy was indeed more pronounced for the high-density than for the low-density panels (cf. Table 1).

At first sight, however, the foregoing analysis may appear paradoxical: increasing informational content by increasing the number of cells increases the perspective effect; at the same time, increasing redundancy, which is equivalent to a decrease in informational content relative to the baseline, also increases the effect. In order to resolve this paradox, it is essential to distinguish between the informational content of the stimulus array, which represents essentially the structural complexity of the array, and the "information to depth" provided by such an array. In order to elucidate this point, let us examine more closely the actual role played by the information variable as manipulated here. Increasing the number of cells filled in, or more particularly the proportion of cells, up to .5, increases the effect, by providing the observer with a greater amount of visual information as regards the progressive deformation of the field from the bottom to the top of the panel.[2] Increasing the regularity of the arrangement of the cells by restricting them to a limited set of columns, while decreasing the information, or introducing redundancy, likewise heightens the effect due to the "overdetermination" of the location of the vanishing point, towards which all the columns converge. These considerations underlie our perhaps somewhat capricious distinction between "amount" and "redundancy" of information, to deal with the density variable and the regularity variable, respectively, as well as our use of the admittedly imprecise term "information to depth."

Speaking more generally, it becomes apparent that any application of an informational model in this area cannot proceed blindly, but must consider the particular ways in which informational content is varied, and their bearing on the perceptual situation and on *S*'s task. Particularly is this true with respect to the role of redundancy, which can probably operate in very different directions, depending on the way it is imposed. For instance, if the restrictions imposed in the selection of cells for the NR fields had involved the rows rather than the columns, the formal amount of redundancy thus introduced would have been the same, yet the effects of this redundancy on the judgments in this task would probably have been less pronounced.

Nevertheless, judiciously applied, the concepts and principles of information theory should prove rewarding in carrying the study of space perception beyond the investigation of isolated cues to the kind of parametric and systematic analysis of the information in the stimulus array on which Gibson has repeatedly insisted. A limited example of this point from our experiment is the treatment of texture density and linear perspective in terms of the more general concept of informational content, which the present approach has made possible.

Summary

This experiment investigated the effects of different stimulus fields, made up of perspective drawings varying in the amount and regularity of the elements subjected to perspective deformation, on the judgment of relative size and distance in the plane of the drawings. Four age groups, varying from first-grade children to college-age adults, were used as *S*s. The results obtained confirmed the prediction that, as the amount and redundancy of information to depth contained in the field increased, the apparent midpoint

[2] For values of $p > .5$, the figure-ground relations would simply be reversed: the information would in effect be concentrated in the white cells, so that the illusion would decrease again until, for $p = 1.0$, the situation would be formally equivalent to that for $p = 0.0$, i.e., to the control panel. This situation is of course faithfully reflected in the concomitant changes in H, the measure of the informational content.

of a segment of a line through the vanishing point would be displaced towards the top of the field. The results for the size judgments were less consistent. The only age difference appeared on the distance judgments, where adults exhibited smaller effects than children between 7 and 14 yr. of age. The implications of the experiment for an informational approach to the study of space perception are briefly considered.

REFERENCES

EDWARDS, A. L. *Experimental design in psychological research.* (Rev. ed.) New York: Rinehart, 1960.

GARDNER, R. W., & LONG, R. I. Errors of the standard and illusion effects with the inverted-T. *Percept. mot. Skills*, 1960, **10**, 47–54.

GIBSON, J. J. *The perception of the visual world.* Boston: Houghton Mifflin, 1950.

GIBSON, J. J. A theory of pictorial perception. *Audiovis. commun. Rev.*, 1954, **1**, 3–23.

GIBSON, J. J. Pictures, perspective, and perception. *Daedalus*, 1960, **89**, 216–227.

GLASSER, O. Optical illusions. In O. Glasser (Ed.), *Medical physics.* Chicago: Yearbook, 1944. Pp. 824–827.

PIAGET, J., & LAMBERCIER, M. Recherches sur le développement des perceptions: III. Le problème de la comparaison visuelle en profondeur (constance de la grandeur) et l'erreur systématique de l'étalon. *Arch. Psychol., Genève*, 1943, **29**, 253–308.

SMITH, O. W. Judgments of size and distance in photographs. *Amer. J. Psychol.*, 1958, **71**, 529–538.

SMITH, O. W., SMITH, P. C., & HUBBARD, D. Perceived distance as a function of the method of representing perspective. *Amer. J. Psychol.*, 1958, **71**, 662–674.

WEINSTEIN, S. The perception of depth in the absence of the texture-gradient. *Amer. J. Psychol.*, 1957, **70**, 611–615.

WOHLWILL, J. F. Developmental studies of perception. *Psychol. Bull.*, 1960, **57**, 249–288.

(Received August 17, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 311-313

STIMULUS GENERALIZATION AS A FUNCTION OF UCS INTENSITY IN EYELID CONDITIONING

JOHN J. PORTER[1]

State University of Iowa

One of the implications of the behavior theory developed by Hull and Spence (Hull, 1943, 1952; Spence, 1956), based on studies of classical and instrumental conditioning, is that generalization performance curves for groups at different drive levels will tend to converge. The derivation of this interaction is as follows: $E_s = H \times D_s$, $E_w = H \times D_w$; then, $E_s - E_w = H(D_s - D_w)$, where D_s = strong drive and D_w = weak drive while E_s and E_w are the corresponding excitatory potentials. If H is the habit strength developed to the training stimulus (S) and $\bar{H}$ that developed to the generalized stimulus (S′), then the above derivation implies that a greater difference in excitatory potential will be expected at S than at S′. Thus an interaction is implied.

Studies of stimulus generalization under two levels of motivation in rats by Brown (1942) and in humans by Rosenbaum (1953) yielded converging generalization curves when these authors' time measures were later transformed to speed measures. Newman (1955) studied stimulus generalization in the rat under different drive levels. Although Newman found evidence of convergence with both speed and extinction measures, the interaction was not statistically significant. Both Jenkins, Pascal, and Walker (1958) and Thomas and King (1959) used pigeons to study generalization at different drive levels in the Skinner box. Jenkins et al. (1958) found divergent relative stimulus generalization gradients under conditions of constant drive differences while Thomas and King (1959) found that three of their four drive groups yielded converging stimulus generalization gradients during extinction.

[1] Now at the University of Wisconsin, Milwaukee.

The author is greatly indebted to Kenneth W. Spence for advice and assistance throughout the course of this investigation.

The present experiment measured the predicted performance convergence of high- and low-drive groups with either a 1500-cps or 400-cps tone after all *S*s were trained at one drive level with the 1500-cps tone alone. An extinction test was used in order to avoid confounding generalization effects with differential reinforcement effects due to differences in puff intensities.

METHOD

Subjects.—The *S*s were 160 women from a course in introductory psychology at the State University of Iowa. Forty-one women were discarded for not meeting a conditioning criterion of more than 8 CRs and less than 36 CRs on Trials 41–80. This conditioning criterion was employed to avoid both ceiling and floor effects when *S*s were shifted to either a higher or lower drive condition during the extinction trials. Six women who gave CRs to the CS alone on test trials were discarded as were 7 others who gave 50% or more responses that met the criterion of voluntary responses used in this laboratory (Spence & Ross, 1959). Four additional *S*s were discarded due to *E* error and 2 due to equipment malfunction. The 100 remaining *S*s were randomly assigned to four groups.

Apparatus and recording method.—The *S*s were seated in an adjustable dental chair in a sound shielded room which was illuminated by a shielded 7.5-w. bulb. The *S*s' room was separated from the recording and control room by a third intervening room. The equipment used to record eyelid responses was the same as that used in previous studies from the Iowa laboratory (cf. Spence & Taylor, 1951). The CS employed during acquisition was the onset of a 1500-cps tone of 70 db. generated by a Hewlett-Packard audio oscillator and delivered by a 6-in. loudspeaker 4 ft. behind *S*. During extinction the same 1500-cps tone served as CS for two groups while the remaining two groups received a 400-cps tone of like intensity. The acquisi-

tion UCS was a 50-msec., .6-psi, air puff delivered to the right eye through a .062-in. diameter orifice by a 110-v. ac solenoid valve. During extinction two groups received a .33-psi puff while the other two received a 2.0-psi puff 2500 msec. after CS onset. A dimly illuminated 2.25-in. diameter circular milk-glass disk located 4 ft. in front of *S* served as a fixation point.

Procedure.—Each *S* was instructed to blink once to the ready signal and then to look at the disk until the tone went off. After the instructions had been read to each *S* she then received three presentations of the CS alone and one presentation of the UCS alone. The intervals between the ready signal and the onset of the CS were 2, 3, or 4 sec. randomly varied. Intertrial intervals of 15, 20, or 25 sec., given according to a fixed schedule, and averaging 20 sec. were used. A CR was recorded whenever the record showed a deflection of 1 mm. or more in the interval 200–500 msec. following CS onset.

Experimental design.—An 80% partial reinforcement schedule was used during training in order to provide for greater resistance to extinction than is obtained with continuous reinforcement. The reinforced and nonreinforced trials followed a prearranged sequence restricted by the provision that no more than 2 nonreinforced trials occurred in each block of 10 trials. On reinforced trials the UCS onset followed the CS by 500 msec.; on nonreinforced trials the UCS began 2500 msec. after the CS onset. McAllister (1953a, 1953b) has shown that little or no conditioning occurs at 2500-msec. intervals, and that the CR extinguishes when the interval is shifted from 500 msec. during acquisition to 2500 msec. during extinction. On both reinforced and nonreinforced trials the CS extended 50 msec. beyond the UCS onset and both CS and UCS terminated simultaneously.

Four groups of 25 *S*s were conditioned and extinguished. All groups received identical (.6 psi, 1500 cps) 80% reinforcement conditions for 80 trials followed by 40 extinction trials under one of four conditions. An extinction test was used in order to avoid confounding of the results through further differential performance build-up under the new puff strengths of 2 psi and .33 psi. Group I received 40 extinction trials with a 2-psi puff and the original 1500-cps tone. Group II differed from I only in receiving a .33-psi puff during extinction. Group III received 40 nonreinforced trials with a 2-psi puff and a 400-cps tone. Group IV differed from III only in receiving a .33-psi puff during extinction.

Results and Discussion

Performance had stabilized at about 60% responding at the point where extinction began. Inspection of the data and an analysis of variance over the last 10 acquisition trials (71–80) revealed no significant performance differences between the four acquisition groups: therefore these four groups were treated as one group.

Three different response measures were used during extinction and three different blocks of trials were analyzed. Extinction results were analyzed in terms of total number of responses, percentage of responses, and superthreshold excitatory potential (E_L) as defined by Spence (1956). The above three measures were taken on Trials 1–10, 1–20, and 1–40, respectively.

Figure 1 presents the percentage of CRs made during Extinction Trials 1–40. The same picture resulted when the other two response measures were used. Examination of Fig. 1 revealed the predicted performance convergence during extinction for the high- and low-drive groups. A similar graphical convergence was obtained on Trials 1–10 and 1–20. Thus there was a smaller difference at the generalized stimulus value than at the original stimulus between the two drive groups. Statistically, such

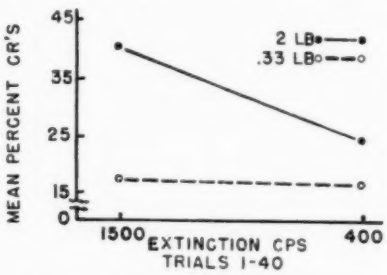


FIG. 1. Mean percentage of conditioned responses for Extinction Trials 1–40.

a condition is represented by a significant interaction between groups. This interaction was significant ($P < .05$, $F = 4.35$ for percentage) over Trials 1–40 for all three performance measures. However this interaction was not significant ($P < .20$) for Trials 1–10 and 1–20 for any of the three performance measures.

Given the interaction over Trials 1–40, the problem becomes one of accounting for the lack of a significant interaction over Trials 1–10 and 1–20 within the framework of the theory used here. As developed earlier, the theory would predict the presence of the sought after interaction within, at the most, a few trials after the change to extinction conditions. Since it required 40 trials to obtain the predicted interaction during extinction it is necessary to examine the effect of inhibition (I) on the predicted interaction. In Hull-Spence theory I is assumed to be a function of the number of nonreinforced trials and is assumed to subtract from the quantity ($D \times H$). If I is added to the two equations used earlier to derive the difference equation it may be seen that I cancels out of the difference equation. This leaves H and D, or $\overline{H}$ and D, as the *differential* factors in this equation. It should be understood that the cancellation of I in the difference equation does not mean that I does not act on performance, but that it does not act differentially. Therefore the interaction predicted by the theory remains unchanged.

Theoretical considerations aside, examination of the range of responses over extinction trial blocks revealed decreasing variability within groups as later extinction trial blocks were examined. This finding suggests that the greater variability of the data during Trials 1–10 and 1–20 may have masked the interaction effect.

Summary

This study investigated the interaction between drive level and original and generalized stimulus conditions during extinction. One hundred *Ss* were first conditioned for 80 trials to respond to a .6-psi air puff on an 80% reinforcement schedule with a CS of 1500 cps. The *Ss* were then divided into four groups which were extinguished with either a 1500- or 400-cps tone and a .33- or 2.0-psi air puff. On all nonreinforced trials during acquisition and extinction the UCS was presented, but 2500 msec. after the CS.

The results confirmed ($P < .05$ for Extinction Trials 1–40) the hypothesized interaction between drive level and original and generalized stimulus conditions predicted by Hull-Spence theory. The effect was not significant over Trials 1–10 or 1–20.

References

Brown, J. S. The generalization of approach responses as a function of stimulus intensity and strength of motivation. *J. comp. Psychol.*, 1942, **33**, 209–226.

Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.

Hull, C. L. A behavior system. New Haven: Yale Univer. Press, 1952.

Jenkins, W. O., Pascal, G. R., & Walker, R. W. Deprivation and generalization. *J. exp. Psychol.*, 1958, **56**, 274–227.

McAllister, W. R. The effect on eyelid conditioning of shifting the CS-UCS interval. *J. exp. Psychol.*, 1953, **45**, 423–428. (a)

McAllister, W. R. Eyelid conditioning as a function of the CS-UCS interval. *J. exp. Psychol.*, 1953, **45**, 417–422. (b)

Newman, J. R. Stimulus generalization of an instrumental response as a function of drive strength. Unpublished doctoral dissertation, University of Illinois, 1955.

Rosenbaum, G. Stimulus generalization as a function of level of experimentally induced anxiety. *J. exp. Psychol.*, 1953, **45**, 35–43.

Spence, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.

Spence, K. W., & Ross, L. E. A methodological study of the form and latency of eyelid responses in conditioning. *J. exp. Psychol.*, 1959, **58**, 376–381.

Spence, K. W., & Taylor, J. A. Anxiety and strength of UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.

Thomas, D. R., & King, R. A. Stimulus generalization as a function of level of motivation. *J. exp. Psychol.*, 1959, **57**, 323–328.

(Received August 18, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 314-317

EXPERIMENTAL EXTINCTION AS A FUNCTION OF NUMBER OF REINFORCEMENTS[1]

JAMES R. ISON

University of Rochester

Overlearning-reversal experiments by Reid (1953), Pubols (1956), and Capaldi and Stevenson (1957) in brightness discrimination and by Pubols (1956) and Ison and Birch (1961) in spatial discrimination have demonstrated facilitation of reversal learning after overlearning compared to control groups reversed at criterion. A similar experiment by Birch, Ison, and Sperling (1960) in differential conditioning demonstrated more rapid extinction of the formerly positive response for the overlearning group and it was concluded that, contrary to the results obtained in the free responding lever press apparatus (Miles, 1956; Perin, 1942; Williams, 1938), resistance to extinction of a running response is not a monotonically increasing function of the number of reinforced trials (N_0). This conclusion was supported, in part, in a runway experiment by North and Stimmel (1960) which demonstrated greater resistance to extinction in a group given 45 reinforcements as compared to groups given 90 or 135. The purpose of the present experiment is to provide further evidence on this relationship in the straight runway apparatus.

METHOD

Subjects.—The *S*s were 75 male hooded rats, approximately 100 days old, obtained from the colony maintained by the Psychology Department of the State University of Iowa. They were randomly assigned to six groups of 12 or 13 *S*s which received either 10, 20, 40, 60, 80, or 100 rewarded acquisition trials.

Apparatus.—A straight alley was housed within a two-unit, black-draped enclosure 4 ft. high, 4 ft. wide, and 11½ ft. long. A plywood panel 4 ft. from the end of the enclosure completely separated the runway section from the goal box except for a hole in the base through which passed the alleyway. The alleyway (covered with glass) was 4 in. high, 3½ in. wide; the start box was 9 in. long, the runway 72 in., and the goal box 18 in. Guillotine retrace doors separated the start box and the goal box from the runway. The entire apparatus was painted flat black and was illuminated by three 75-w. Lumline lights attached to the ceiling of the enclosure directly over the goal box and at distances of 1 and 4 ft. from the start box. These bulbs were screened to give an incident light of approximately 3 ft-c (range, 2.9 to 3.3) on the alleyway. Infrared photobeams permitted the measurement of running speed over two 1-ft. segments of the alleyway beginning 1 and 4 ft. from the start box. The apertures for the lights were ¼ in. in diameter and covered with painted plastic, thus provided little differential stimulation. Mercury switches in the start-box and goal-box doors permitted the measurement of running speed over the entire alleyway. All times were recorded on electronic clocks and the operation of the timing circuits was silent.

Procedure.—The *S*s were allowed 10 gm. of food powder each day in wet mash, presented ½ hr. after each daily treatment. They were placed on this schedule and handled for 2 min. each day for 14 days prior to acquisition training. Two trials were given on Days 1 and 2 and 5 trials per day thereafter except for Extinction Day 1, which contained the final rewarded trial followed by 5 nonrewarded trials. The reward was 0.4 gm. of food powder in 0.3 ml. of water, given in a glass dish 4 in. from the goal-box end wall. In extinction the empty glass dish was present, and *S* was detained in the goal box for 30 sec. The minimum intertrial interval in both acquisition and extinction was 18 min. during which *S* was detained in a wooden and

[1] This research was performed while the author was a Rackham Postdoctoral Fellow at the State University of Iowa. The author wishes to thank K. W. Spence for his generous assistance.

wire mesh box with water available. Extinction was carried to a minimum of 80 trials, continuing if necessary until *S* took longer than 120 sec. to enter the goal box.

Results

Four measures are reported, the number of trials to various extinction criteria, running speed in early extinction, the number of avoidance responses made in extinction, and the number of trials to the first avoidance response.

Trials to criterion.—A criterion trial was the first trial on which *S* exceeded a criterion number of seconds to enter the goal box after the start-box door was opened. Four criteria were chosen, 10, 20, 40, and 120 sec., and the mean numbers of trials to each of these are presented in Fig. 1. With the criteria of 10 and 20 sec., the groups did not differ ($F < 1.00$). With the 40-sec. criterion the difference among groups was significant at the .01 level ($F = 3.69$, $df = 5/69$) and on the 120-sec. criterion the difference was significant at the .001 level ($F = 7.48$, $df = 5/69$). On both of these latter criteria the relationship between N_g and trials to criterion was negative.

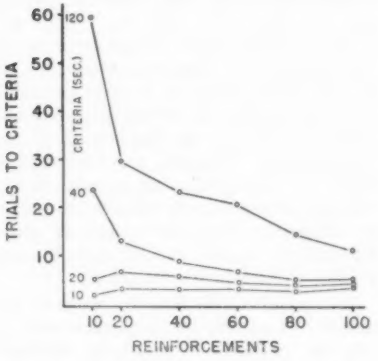


Fig. 1. The mean number of trials to criteria of 10, 20, 40, and 120 sec. between *E*'s opening the start box and *S*'s entering the goal box.

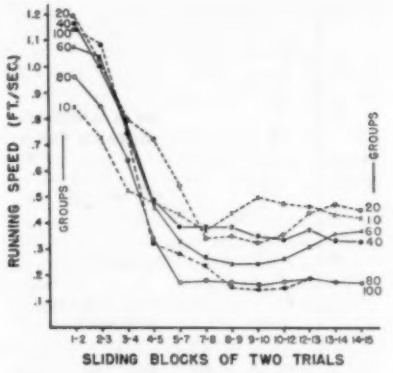


Fig. 2. Mean running speed in sliding blocks of two trials over the first 3 days of extinction. (The first trial of each day is omitted.)

Running speed in extinction.—In Fig. 2 is depicted the running speed (over the entire alleyway) of the six groups in the first 3 days of extinction. The other speed measures showed essentially identical results. On the initial point Groups 20, 40, 60, and 100 were but little different whereas Groups 10 and 80 were slower. In subsequent trials Group 100 showed the greatest decrement and, with Group 80, was responding at the slowest speeds at the end of the 3 days. Group 10 decreased the least and the other three groups fell roughly in order between Groups 10 and 80. The groups, ranked in mean response speed per trial, were Group 20 ($M = .60$ fps); Group 40, ($M = .55$ fps); Group 10, ($M = .52$ fps); Group 60, ($M = .50$ fps); Group 100, (M = .41 fps); Group 80, ($M = .37$ fps). A trial by trial mixed analysis of variance (Lindquist, 1953) yielded a significant Groups effect ($F = 3.64$, $df = 5/69$, $P < .01$), a significant Trials effect ($F = 77.77$, $df = 12/828$,

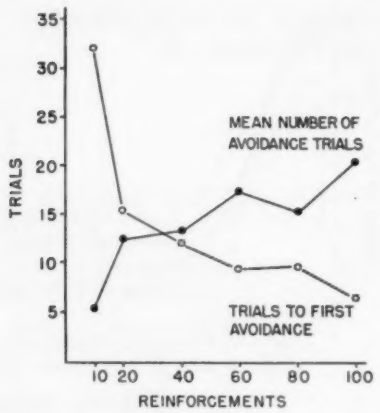


FIG. 3. The mean total number of extinction trials on which an avoidance response occurred and the mean number of trials to the first avoidance response.

$P < .001$), and a significant Groups × Trials interaction ($F = 1.96$, $df = 60/828$, $P < .01$).

Avoidance responses in extinction.— An avoidance response was recorded whenever *S* turned and moved in the direction of the start box. In Fig. 3 two different measures are shown. One is the mean number of trials before the first avoidance response occurred. A simple analysis of variance of these data yielded a significant Groups effect ($F = 4.10$, $df = 5/69$, $P < .005$); the relationship between this variable and N_g was negative. The second is the mean number of trials on which an avoidance response occurred in the 80 extinction trials. An analysis of these data yielded a significant Groups effect, ($F = 5.14$, $df = 5/68$[2], $P < .001$); the relationship between this variable and N_g was positive.

[2] One *S* was dropped after Extinction Trial 30 because of an experimental error. This reduced the *df* for this comparison to 68.

DISCUSSION

Under the conditions of this experiment, trials to extinction criteria of 40 and 120 sec. were negatively related to N_g and running speed in the first 3 days of extinction was nonmonotonically related to N_g. These data support the conclusions of Birch, Ison, and Sperling (1960) in their account of overlearning-reversal problems and confirm and extend the findings of North and Stimmel (1960).

This relationship between trials to criterion and N_g is to be contrasted with the negatively accelerated increasing function typically obtained in the Skinner box. One possible reason for this difference is that in the present experiment and in that of North and Stimmel (1960) the reward magnitude (W_g) was large; whereas, in the Skinner box the reward was relatively small. Several experiments have suggested that N_g and W_g interact in determining resistance to extinction, the relationship between R_n and W_g being positive at small N_g (Zeaman, 1949) but negative at large N_g (Armus, 1959). This interaction might be reversible, i.e., at large W_g, R_n is negatively related to N_g; whereas, at small W_g, R_n is positively related to N_g over at least part of the range. A second possibility is that the form of the function is determined by some characteristic of the investigated response. The one overlearning-reversal experiment which did not use a running response (McCulloch & Pratt, 1934) found that extended training retarded the subsequent reversal, which is contrary to the results of the later studies. Whether nonmonotone or negative relationships are peculiar to the running response and whether they can be obtained with other responses given appropriate values of N_g and W_g are subject to further investigation.

The positive relationship between the number of avoidance responses and N_g and the negative relationship between trials to the first avoidance response and N_g are consistent with interference theories of extinction stressing the acquisition of frustration-instigated avoid-

ance responses which compete with the approach response (e.g., Birch, 1961; North & Stimmel, 1960). Following Amsel (1958) and Spence (1960), the magnitude of frustration elicited on nonreinforced trials is assumed to be in part a positive function of the number of prior reinforcements, which, with the further assumption that the strength of the avoidance response is positively related to frustration magnitude, is sufficient to account for these two relationships.

Summary

The relationship between the resistance to extinction of a running response and the number of acquisition trials (N_a) was investigated. Six groups of rats received either 10, 20, 40, 60, 80, or 100 rewarded trials followed by 80 nonrewarded extinction trials at five trials per day with an intertrial interval of 18 min. The mean numbers of trials to extinction criteria of 40 and 120 sec. were negatively related to N_a, and running speed in early extinction was nonmonotonically related to N_a. These data were contrasted with those previously obtained in the Skinner box. In addition, the mean number of trials on which avoidance responses occurred in extinction was positively related to N_a and the mean number of trials to the first avoidance response was negatively related to N_a. These latter relationships are consistent with interference theories of extinction which stress the acquisition of competing avoidance responses.

References

Amsel, A. The role of frustrative nonreward in noncontinuous reward situations. *Psychol. Bull.*, 1958, **55,** 102–119.

Armus, H. L. Effect of magnitude of reinforcement on acquisition and extinction of a running response. *J. exp. Psychol.*, 1959, **58,** 61–63.

Birch, D. A motivational interpretation of extinction. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1961.* Lincoln: Univer. Nebraska Press, 1961. Pp. 179–196.

Birch, D., Ison, J. R., & Sperling, S. E. Reversal learning under single stimulus presentation. *J. exp. Psychol.*, 1960, **60,** 36–40.

Capaldi, E. J., & Stevenson, H. W. Response reversal following different amounts of training. *J. comp. physiol. Psychol.*, 1957, **50,** 195–198.

Ison, J. R., & Birch, D. T maze reversal following differential endbox placement. *J. exp. Psychol.*, 1961, **62,** 200–202.

Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.

McCulloch, T. L., & Pratt, J. G. A study of the presolution period in weight discrimination by white rats. *J. comp. Psychol.*, 1934, **18,** 271–290.

Miles, R. C. The relative effectiveness of secondary reinforcers throughout deprivation and habit strength parameters. *J. comp. physiol. Psychol.*, 1956, **49,** 126–130.

North, A. M., & Stimmel, D. T. Extinction of an instrumental response following a large number of reinforcements. *Psychol. Rep.*, 1960, **6,** 227–234.

Perin, C. T. Behavior potentiality as a joint function of the amount of training and degree of hunger at the time of extinction. *J. exp. Psychol.*, 1942, **30,** 93–113.

Pubols, B. H. The facilitation of visual and spatial discrimination by overlearning. *J. comp. physiol. Psychol.*, 1956, **49,** 243–248.

Reid, L. S. The development of noncontinuity behavior through continuity learning. *J. exp. Psychol.*, 1953, **46,** 107–112.

Spence, K. W. *Behavior theory and learning.* Englewood Cliffs, N. J.: Prentice Hall, 1960.

Williams, S. B. Resistance to extinction as a function of the number of reinforcements. *J. exp. Psychol.*, 1938, **23,** 506–522.

Zeaman, D. Response latency as a function of amount of reinforcement. *J. exp. Psychol.*, 1949, **39,** 466–483.

(Received August 21, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 318-324

EYE-MOVEMENT LATENCY, DURATION, AND RESPONSE TIME AS A FUNCTION OF ANGULAR DISPLACEMENT [1]

ALBERT E. BARTZ [2]

Applied Research Laboratory, University of Arizona

Characteristics of eye movements and reaction time to visual stimuli have both been studied systematically since the beginning of the twentieth century. Various components of the visual response were systematically investigated by Dodge and others (Diefendorf & Dodge, 1908; Dodge & Cline, 1901). Using photographic recording techniques, they found that the average latency was about 200 msec., while the eye movement duration was 29 msec. for a 5° movement, and increased to 100 msec. for a 40° movement. Essentially the same results were obtained by Miles (1936) and Hackman (1940).

It must be noted that these time intervals do not reflect the time involved in the process of "seeing" an object in the periphery. After the eye has fixated upon the peripheral stimulus, the observer still must process the new information and make some response.

More recently other investigators have been concerned with the total response time (RT) when there is more than a simple movement involved. Hyman (1953) found that the total RT increased when the task required *S* to identify the specific location of the stimulus. Words were assigned to various lights and the RT was measured by a voice key set off when *S* pronounced the correct word for the stimulus location. It should be noted that this increase in RT occurred even though *S* was not specifically instructed to move his eyes since Hyman's stimulus lights subtended a maximum of only 2.5° of visual angle. This type of RT is more closely related to the problem of seeing, since the total visual reaction must include an identification of what is seen. As was expected Hyman found this identification type of RT to be longer—the lengthening being a function of the statistical probability that a stimulus would appear in the specific location identified. Hyman's vocal RTs varied from 300 to 750 msec.

However, this complex response still does not represent accurately the process of "seeing" an object in the periphery. To see an object in the periphery *S* must not only identify the location and swing his eyes to it, but also must interpret the stimulus. The present research was designed for two purposes. Experiment I involved the investigation of RTs as functions of angular displacement from the line of regard and the number of stimuli to which *S* must attend. Experiment II was designed to isolate and measure the various components of the total RT. By using the electrical

[1] This paper is based upon a thesis submitted to the Graduate School of the University of Arizona in partial fulfillment of the requirements for the PhD degree. The author is indebted to his advisor, Neil R. Bartlett, for helpful criticism and encouragement. Gratitude is also expressed to the Highway Research Board for Contract AFE 8093 that made the project possible, and to R. W. Lansing, R. L. Lucas, and H. Tucker for their criticism and suggestions.

[2] Now at Concordia College, Moorhead, Minnesota.

method for recording eye movements (Ford & Leonard, 1958; Mowrer, Ruch, & Miller, 1936), it was possible to isolate the latency, eye-movement duration, and the time required for interpreting the stimulus.

Experiment I

Method

Subjects.—The *S*s were 3 20-yr.-old male volunteer undergraduates. They were free from pertinent visual defects as measured by an Orthorater.

Apparatus.—The peripheral stimuli were arranged in a semicircle about *S*, and were at 40°, 20°, 10°, 5°, and 2.5° right and left. There was also one stimulus at the center or 0°. A point between *S*'s eyes was the center of a circle 6 ft. in radius, and the 11 stimuli were at eye level.

The stimuli were the digits 4, 5, 6, and 9 presented by Burroughs Type BD200S Nixie indicator tubes. The height of each numeral was .305 in., subtending a visual angle of 14′ at a distance of 6 ft. These four digits were chosen on the basis of preliminary tests as giving a good response for triggering the voice relay circuit.

The indicator tubes were mounted on a curved panel painted a flat gray to minimize glare. The only illumination in the room was a fluorescent source located 6 ft. behind and 3 ft. above *S*, giving an illumination on the panel of 2.78 mL.

To keep *S* looking at the center of the display prior to the presentation of a stimulus, a tracking task requiring continuous monitoring was used. This task required *S* to follow a light moving in a triangular pattern with another light controlled by a three-button switch. The triangle subtended 1.5° at *S*'s eye, and *S* could perform satisfactorily while fixated on the center of the triangle.

The *S* sat in an armchair with a headrest, preventing any horizontal head movements, and wore a headset with an attached carbon microphone. The microphone triggered the voice relay circuit whenever *S* responded verbally to the number presented in any one of the indicator tubes. This circuit tripped a latching relay that stopped the timer.

The *E* had controls to select an indicator light at any of the 11 positions and any of the 4 numerals appearing in it.

Procedure.—In both the training and experimental sessions, the procedure for the presentation of the stimuli was the same. The *S* entered the experimental room which was light-proofed to exclude any extraneous illumination that might reflect from the curved surfaces of the indicator tubes. After the headrest and microphone were adjusted, 10 warm-up trials were given before the session began.

At the start of each trial *S* began tracking the center display of lights. At intervals of 3, 4.5, or 6 sec. after the start of the tracking task, a number in one of the indicator tubes came on and the tracking lights extinguished. The *S* then moved his eyes to the position of the stimulus and verbalized the number into the microphone. This response stopped the timer and extinguished the number in the indicator tube. After a 5-sec. rest period the next trial was begun, and *S* resumed his tracking task. There were two 1-min. rest periods during the experimental session.

Experimental design.—To insure reliability of results, all *S*s were highly trained prior to the beginning of the experimental sessions. For the training trials the indicator lights in the 20° right and left positions were used. Each *S* made 144 responses per session, 72 to each position. The four numerals appearing in the indicator lights were randomized among the 72 stimuli for each position. The training sessions were concluded when both the means and *SD*s became asymptotic. This occurred on Day 16 for 1 *S* and on Day 19 for the other 2. After the training trials were concluded each *S* experienced two sessions of responding to all 11 lights.

The experimental trials were run for 12 days and were initiated on the day following the training sessions. In order to determine how RT varies as a function of the number of possible stimuli, it was necessary to divide Exp. I into Sequences A and B. The stimuli in Sequence A consisted of the indicator lights at the 20° and 10° right and left positions. Each of the 16 possible combinations of position and indicator numeral appeared three times in each group of 48 trials, and was randomized throughout each group. The entire session of 144 trials consisted of three of these groups of 48. Sequence A was presented on Experimental Days 1, 2, 11, and 12.

The stimuli in Sequence B consisted of the indicator lights in all 11 positions. To simplify the data reduction the light at the center position was considered as two stimuli, with one-half of the responses counting on the left side of the visual field and the other half counting on the right side. As a result there were 48 possible combinations of position and indicator numeral. Each of the combinations

TABLE 1

MEAN RTs AND *SD*s (.01 SEC.) TO STIMULUS POSITIONS IN SEQUENCE B: EXP. I

Position	Direction			
	Left		Right	
	Mean	*SD*	Mean	*SD*
40°	91.12	8.3	90.05	7.8
20°	77.70	7.2	77.14	6.2
10°	72.82	7.4	72.13	5.9
5°	70.35	7.4	69.99	7.3
2.5°	67.79	7.8	66.81	7.7
0°	58.43	6.9	—	—

appeared once in each group of 48 trials and was randomized throughout. Again, each experimental session consisted of three groups of 48 trials. Sequence B was presented on Experimental Days 3 through 10.

Results

The data from Sequence B, with all 11 positions, were analyzed by means of a four-factor (Position, Numeral, *S*, and Day) analysis of variance. All of the main factors were significant beyond the .001 level. Mean RTs to the lights at various positions in the periphery and their *SD*s are given in Table 1.

The four numerals, or the four vocal responses required (4, 5, 6, and 9), yielded significantly different mean RTs. An examination of the means showed that the vocal responses 4 and 5 were significantly faster than the response 6 or 9. The significant Days effect was due to lengthened RTs occurring on Days 3 and 4.

The Position × Numeral interaction was significant at the .001 level. This indicated that at some positions certain digits yielded faster RTs than at other positions. An inspection of the means showed that in 8 of the 11 positions the fastest mean response was made to the numeral 4. However, in the other three positions, 40°, 20°, and 10° left, the fastest response was made to the numeral 5.

The significant Position × *S* interaction indicated that at some positions certain *S*s performed better than others. Inspection of the means showed that 1 *S* was faster at the 40° and 20° positions and slower on the other positions.

The Numeral × *S* interaction indicated that some *S*s responded faster to certain numerals.

The significant Position × Numeral × Day and the Numeral × *S* × Day interactions indicated that the significant Position × Numeral and Numeral × *S* interactions varied as a function of the day on which the responses were made. The results of the analysis of variance are shown in Table 2.

The data from Sequence A, with only four positions used (20° and 10° right and left), were analyzed by a similar analysis of variance. The results of this analysis were identical with that of Sequence B, with the exception of the main effect of Day. There was no significant difference be-

TABLE 2

ANALYSIS OF VARIANCE OF RTs: EXP. I

Source	*df*	*MS*	*F*
Position (P)	11	10,137.64	1,062.65**
Numeral (N)	3	1,891.67	198.29**
S	2	8,683.35	910.20**
Day (D)	7	113.29	11.88**
P×N	33	34.21	3.59**
P×*S*	22	64.23	6.73**
P×D	77	10.71	1.12
N×*S*	6	138.83	14.55**
N×D	21	11.95	1.25
S×D	14	61.36	6.43**
P×N×*S*	66	15.86	1.66*
P×N×D	231	7.78	—
P×*S*×D	154	10.33	1.08
N×*S*×D	42	16.07	1.68*
P×N×*S*×D	462	9.54	
Total	1,151		

* $P < .05$.
** $P < .001$.

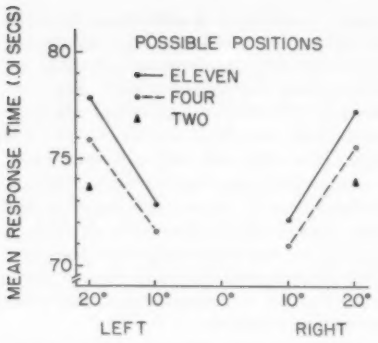


FIG. 1. Response time as a function of number of possible stimuli.

tween responses made on the 4 days of Sequence A, i.e., Days 1, 2, 11, and 12. This also indicated that the counterbalancing technique used to offset possible transfer effects was successful.

The use of these two sequences was to enable a comparison of responses made to 11 possible positions (Sequence B) as against 4 possible positions (Sequence A). Shown in Fig. 1 is the comparison of the responses at the 20° and 10° left and right positions from both sequences. As is evident from the graph, the mean RTs were faster for the smaller number of possible stimulus positions. All differences were significant at the .05 level. Also shown in Fig. 1 are the data from a situation in which only the 20° left and right positions were used. These means were taken from the last 5 days of the training trials.

EXPERIMENT II

Method

Apparatus.—The same apparatus was used for presenting the stimuli as in Exp. I. To record the eye movements necessary for measuring the components of the total response, electrodes were placed behind the external canthi of *S*'s eyes. The output from the electrodes was fed to a Grass Model P-5 preamplifier, and the output of the preamplifier terminated at an oscilloscope. The upper trace of this dual channel oscilloscope was a record of *S*'s eye movements. For the lower trace the input was from the first stage of amplification of the electronic voice key. The sweep was triggered when an indicator light came on. A Dumont oscilloscope camera was used to photograph the tracings.

Procedure.—Experiment II was begun on the day following the close of Exp. I. As before, *S* attended to the tracking task until an indicator light came on, moved his eyes to the stimulus, and responded verbally to the indicator numeral.

Experimental design.—Because of the time required to manipulate the camera for recording the CRT trace, it was necessary to reduce the number of responses during the experimental session. It was also necessary to alter the order of presentation of the indicator lights. The order was arranged so that only every third response was recorded by the oscilloscope camera. Since the main interest was in responses to stimuli involving eye movement, only those responses to the 40°, 20°, 10°, and 5° right and left positions were recorded. To insure high reliability in the vocal response, all responses recorded were to the stimulus 5. The remaining three numerals were divided equally among the unrecorded stimuli.

Each *S* made a total of 96 responses at each session. Of these 32 were recorded, 4 for each of the eight positions. The *S*s were not aware that only some responses were being recorded. After 4 days of testing, *S*s made

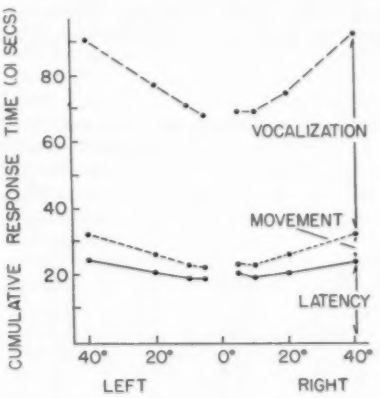


FIG. 2. Portion of response occupied by three visual components.

a total of 384 recorded responses, or 48 for each of the eight positions.

Results

The proportions of the total RTs accounted for by the three components of latency, movement, and vocalization are shown in Fig. 2. As was expected, both eye-movement latency and eye-movement duration increased as the angle from the line of regard increased. However, vocalization time (the interval between the cessation of the eye movement and *S*'s vocal response) also increased with angle. The means and *SD*s for latency, duration, and vocalization are shown in Table 3.

No significant practice effect appeared during the 4 experimental days. As in Exp. I there were significant subject differences, appearing mostly in the latency and vocalization.

Discussion

The type of RT that was investigated in this experiment involved much more than a simple eye-movement latency, so it was logical to expect greater RTs than the latencies reported by Dodge, Miles, or Hackman. Their interest was not in the speed of "seeing," but only in the time that was required for the eyes to begin moving to a peripheral stimulus. In Exp. II it was found that the average eye-movement latency agreed quite well with these previous studies. The overall mean latency to stimuli at all angles was .213 sec. (SD = .041 sec.). This coincides very well with previous data.

However, as mentioned earlier, data on the latency of the ocular reaction does not accurately reflect the process required to see objects in the periphery. The *S* must get his eyes in motion, swing his eyes to the new object, and then make his response.

It can be seen from Fig. 2 that it is inaccurate to state an "average" value for eye-movement latency, since the time required for the eyes to begin their movement was a function of the angle at which the new stimulus was located. The fact that eye-movement latency increases as the angle from the center line of regard increases was noted earlier by White, Eason, and Bartlett (1962).

It is further apparent that RTs must increase as a function of the angle through which the eyes must move. The actual movement of the eye takes longer as the angle from the center line of regard increases. However, this time interval was extremely small, accounting for only 5% to 10% of the total RT.

A close inspection of the total RT as a function of position shown in Table 1 yielded an interesting observation. The large differences between mean RTs at

TABLE 3

Means and *SD*s for Eye-Movement Latency, Duration, Vocalization, and Total Response in Exp. II

Position	Latency		Duration		Vocalization		Total	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
40° L	24.8	5.3	8.7	1.0	56.0	5.9	89.5	8.5
20° L	21.2	4.4	5.7	0.8	50.5	5.9	77.4	8.0
10° L	19.6	2.4	3.8	0.6	47.9	5.8	71.5	9.5
5° L	19.7	2.5	3.1	0.5	45.6	5.7	68.5	6.5
5° R	20.8	2.5	3.2	0.5	45.8	6.1	69.8	6.8
10° R	19.5	2.5	3.8	0.6	46.6	6.1	69.9	6.7
20° R	20.7	3.9	5.7	0.9	49.0	6.7	75.3	6.9
40° R	24.2	4.2	8.7	1.2	57.5	8.1	90.4	8.7

the various positions could not be fully explained by differences in eye-movement latency or duration. For example, the mean RT to the 40° left position was .911 sec. and that of the 20° left position was .777 sec., for a difference of .134 sec. The difference in eye-movement latency to the two positions was .036 sec., and the difference in eye-movement duration was only .030 sec. Thus latency and movement accounted for less than half of the original difference between the total RTs to the two positions.

Since two of the three components have been accounted for, the difference must be due to the third component, the vocalization time. As defined earlier, vocalization was the time interval between the completion of the eye movement and *S*'s vocal response. As can be seen from Table 3, the time required to make this response after *S* was looking at the signal varied as a function of indicator light position. In terms of the experimental situation, it took longer for *S* to "recognize" the numeral presented and verbalize the response when the stimulus was at a greater angle in the periphery.

There are several possible explanations that may be advanced to account for this observation. With the type of recording method used, it was difficult to distinguish very small eye excursions in comparison to the gross movement at the onset of the signal. Therefore, if at the end of a large angular movement, the eye hunts for an exact fixation, this hunting would probably occupy several degrees or less. Small hunting excursions could not be read from the type of records taken. It is possible that this hunting time occupies proportionately more time as the angle increases. This hunting may be comparable to the variable error discussed by Woodworth (1899) in simple motor tasks. He found that variable error for controlled rapid movements increased with amplitude or distance traveled. Fitts (1954) has suggested that in the concept of fixed information-transmission capacity of the motor system, such increased variability is due to decreased information that the movement provides. It is possible that the hunting may be the variable error of a simple motor response.

Another possible explanation is in terms of accommodation. As the eyes move, the muscles exert different patterns of tension upon the two eyes, and therefore some accommodation may be required in fixating upon a peripheral stimulus. Because of binocular convergence this effect is probably magnified for stimuli close to the eyes. Although the stimuli were 6 ft. from *S*, sufficient accommodation may be required to make the effect appreciable.

In an experiment mentioned previously (Hyman, 1953), it was reported that RT increased as a function of the probability that a stimulus would appear in a specific location. It would follow from this evidence that RT in general will increase with an increase in the number of possible stimuli to which *S* must react. As was noted in Fig. 1 the RTs for 11 possible stimuli were significantly longer than for 4 possible stimuli. (A decrease was also noted in the training trials when 2 possible stimuli were used.)

In the 8 days with 11 stimuli, there appeared to be no further learning taking place. With the Days main effect of Sequence A nonsignificant, it can be assumed that no learning in that situation took place over a 12-day period. However, considering the data of Mowbray and Rhoades (1959), it is entirely possible that RTs could be significantly shortened with practice over a long period of time.

Although some investigators (Hackman, 1940) have found shorter RTs to stimuli on the right side, there were no significant differences in this experiment between RTs to stimuli on the left and right sides.

In the general use of the term, the process of "seeing" refers to a wide variety of functions. The usual figure given for RTs to a visual signal is about one-fifth of a second. In such situations the practiced *S* makes some kind of manual response to the onset of a visual signal. However, a review of the literature shows a wide range of values for

RT—obviously a function of the experimental conditions. In this study, requiring more than a simple reaction, the range in RTs was from .584 sec. for the central stimulus to .906 sec. for extreme positions, with an overall mean of .727 sec.

Summary

The present research was initiated with two purposes in mind: (*a*) to determine the speed of seeing in a complex visual task (Exp. I), and (*b*) to isolate and measure the various components of the total response (initial latency, travel time of the eye, and the response time for interpreting the signal).

Results of Exp. I showed that RT increased as the angle from the center line of regard increased. There was no significant difference between pairs of means for right and left sides. It was also found that response time increased as the number of possibie signals increased. In Exp. II, the time required for each of the three components of the response increased as the angle increased. Several interpretations of the positive relationship between angle and the time required for *S* to make his vocal response after his eyes had reached the signal were considered.

References

Diefendorf, A. R., & Dodge, R. An experimental study of the ocular reactions of the insane from photographic records. *Brain*, 1908, **31**, 451–489.

Dodge, R., & Cline, T. S. The angle velocity of eye movements. *Psychol. Rev.*, 1901, **8**, 145–157.

Fitts, P. M. The information capacity of the human motor system in controlling the amplitude of movement. *J. exp. Psychol.*, 1954, **47**, 381–391.

Ford, A., & Leonard, J. L. Techniques for recording surface bioelectric direct currents. *USN Electron. Lab. res. Rep.*, 1958, No. 839.

Hackman, R. B. An experimental study of variability in ocular latency. *J. exp. Psychol.*, 1940, **27**, 546–558.

Hyman, R. Stimulus information as a determinant of reaction time. *J. exp. Psychol.*, 1953, **45**, 188–196.

Miles, W. R. The reaction time of the eye. *Psychol. Monogr.*, 1936, **47**(2, Whole No. 212).

Mowbray, G. H., & Rhoades, M. V. On the reduction of choice reaction times with practice. *Quart. J. exp. Psychol.*, 1959, **11**, 16–23.

Mowrer, O. H., Ruch, T. C., & Miller, N. E. The corneo-retinal potential difference as the basis of the galvanometric method of recording eye movements. *Amer. J. Psychol.*, 1936, **114**, 423–428.

White, C. T., Eason, R. G., & Bartlett, N. R. Latency and duration of eye movements in the horizontal plane. *J. Opt. Soc. Amer.*, 1962, **52**, 210–213.

Woodworth, R. S. The accuracy of voluntary movement. *Psychol. Monogr.*, 1899, **3**(2, Whole No. 13).

(Received August 24, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 325-326

SUPPLEMENTARY REPORT: DIRECTION OF CHANGE IN CS IN EYELID CONDITIONING[1]

FRANK A. LOGAN AND ALLAN R. WAGNER

Yale University

The assumption that the important parameter of the CS is the amount of change from the pre-CS condition to the CS condition (e.g., Logan, 1954; Perkins, 1953) implies that a decrease in intensity should be as effective a CS as the corresponding increase in intensity. The assumption that the absolute value of the CS has a motivational (dynamogenic) property (e.g., Hull, 1952) implies that an increase in intensity should be more effective. Kish (1955) found tone-off to be a less effective CS than tone-on for avoidance conditioning in rats but Schwartz and Goodson (1958), using a comparable situation, found these events to be equally effective. Hansche and Grant (1960) concluded that light-off was as effective as light-on for eyelid conditioning under a procedure in which the light was off between trials for all *Ss*. The present study compares an increase with a decrease in intensity between two nonzero values treated symmetrically.

Method.—The general features of the eyelid conditioning apparatus, recording equipment, and procedures have been described elsewhere (Dufort & Kimble, 1958). The CS was provided by a circular milk glass disk, 2.25 in. in diameter, set in a flat black ground, and illuminated from behind by General Electric NE30 neon bulbs. The onset of the CS was either an increase from two to four bulbs or a decrease from four to two bulbs. In each case, the CS intensity lasted for 600 msec. during the last 100 msec. of which a 2-lb. air-puff CS was delivered to the corner of the eye. The non-CS intensity remained on during the intertrial interval which averaged 20 sec. in length.

Five test trials of CS or UCS alone were followed by 60 conditioning trials. During these trials the CS for half of the *Ss* was an increase while for the other half it was a decrease in illumination. All *Ss* were then given 20 additional conditioning trials with the opposite CS. The results from 16 female student nurses were combined with those from 40 male undergraduates since they were virtually identical.

Results and discussion.—The results are shown in Fig. 1. Both the increase and decrease in intensity were clearly and equally effective CSs in producing a relatively high level of conditioning. Although the null hypothesis cannot be proven statistically, the standard error of the difference between the groups at the end of training was only 7% and hence it is unlikely that the true difference deviates very much from zero. The data thus indicate the greater relative importance of the change parameter of the CS rather than its absolute intensity.

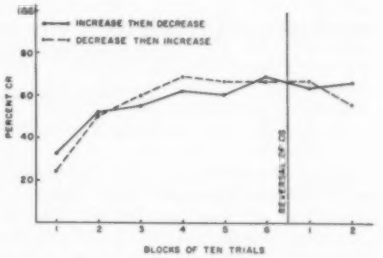


FIG. 1. Percentages of conditioned responses during training. (The solid curve refers to *Ss* for whom the CS was an increase from the between-trials intensity during the first 60 trials and a decrease during the next 20 trials, while the dashed curve refers to *Ss* who received these CS conditions in the reverse order.)

The degree of transfer when the direction of change was reversed is remarkable. Indeed, a slight drop in performance would be expected because of the "extinction trial" given inadvertently when the non-CS intensity was reversed between the last acquisition trial and the first reversal trial. This finding suggests that generalization should be viewed in terms of a surface including the non-CS condition as well as the CS condition. However, it will require a large parametric study adequately to characterize this surface.

[1] Supported in part by Grants G-9014 and G-13080 from the National Science Foundation.

REFERENCES

DUFORT, R. H., & KIMBLE, G. A. Ready signals and the effect of interpolated UCS presentations in eyelid conditioning. *J. exp. Psychol.*, 1958, **56**, 1-7.

HANSCHE, W. J., & GRANT, D. A. Onset versus termination of a stimulus as the CS in eyelid conditioning. *J. exp. Psychol.*, 1960, **59**, 19-26.

HULL, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.

KISH, G. B. Avoidance learning to the onset and cessation of conditioned stimulus energy. *J. exp. Psychol.*, 1955, **50**, 31-38.

LOGAN, F. A. A note on stimulus intensity dynamism (*V*). *Psychol. Rev.*, 1954, **61**, 77–80.
PERKINS, C. C., JR. The relation between conditioned stimulus intensity and response strength. *J. exp. Psychol.*, 1953, **46**, 225–231.
SCHWARTZ, M., & GOODSON, J. E. Direction and rate of conditioned stimulus change in avoidance performance. *Psychol. Rep.*, 1958, **4**, 499–502.

(Received June 30, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 326–327

SUPPLEMENTARY REPORT: THE EFFECT OF STIMULUS DURATION AND LUMINANCE ON VISUAL REACTION TIME [1]

DAVID RAAB AND ELIZABETH FEHRER

Brooklyn College

Raab, Fehrer, and Hershenson (1961) found that simple reaction time (RT) was independent of stimulus duration over the range of 10 to 500 msec. Luminance, on the other hand, was found to be an important determiner of RT. Since intensity rather than total energy (intensity times duration) determined RT, it is obvious that the critical duration (CD) for RT is 10 msec. or less for the three luminance levels (3000, 30, and 0.3 ft-L) investigated.

The term critical duration has been borrowed from visual threshold studies, which have shown reciprocity (Bunsen-Roscoe law) up to a CD of approximately 100 msec., beyond which temporal integration ceases and the threshold is defined solely in terms of luminance.

It seemed worthwhile to determine the CDs in the mediation of RT for the luminances previously studied and the relation between RT and stimulus duration below these critical values. In the experiment to be reported, the six durations ranged from 0.5 to 20 msec., and thus overlapped the range used previously. Two additional intermediate luminances were included.

Method.—Target flashes were generated and RTs measured by the same equipment as that employed in our previous study. A single Tektronix wave-form generator provided the gating pulses for the glow modulator tube; pulse durations were switched between trials, as required.

In order to generate flashes having wave forms as rectangular as possible, the driving pulses were shaped to "overvolt" the glow modulator tube, and the tube itself was placed next to an ultraviolet source. With these arrangements, flash energy was found to be proportional to flash duration within 0.5 db. from 0.5 to 20 msec. The circular target, 1 cm. in diameter, subtended 1° 10′ of arc and was viewed binocularly.

Two senior honors students and the 2 authors served as *S*s. Each *S* served in 30 experimental sessions. Computations are based on data of the last 25 sessions. Only one luminance was used in a given session; the five luminances were counterbalanced over test days for each *S*. Each session consisted of four blocks of 18 trials each, in which each combination of the six durations and three foreperiods appeared once in random order. Only the four longer durations could be explored for the 0.3 ft-L luminance, since this light was below foveal threshold when presented for 0.5 or 1 msec.

Each session began with 5 min. of dark adaptation. Four practice trials preceded the recorded trials. The four blocks were separated by 1-min. rest periods.

The 12 RTs obtained in a session for each of the six durations were reduced to 10 by

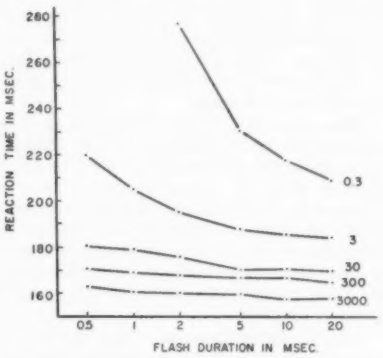


FIG. 1. Reaction time as a function of stimulus duration. (The parameter is flash luminance in ft-L. Each data point is the mean for 4 *S*s.)

[1] This research was supported by grants from the National Science Foundation (G-6456) and from the National Institute of Neurological Diseases and Blindness (B-1028) and by funds provided by Brooklyn College. The data were gathered by Carlos Goldberg and Naomi Maizel as part of an honors course.

discarding the longest and the shortest RT. Testing over 25 days (5 at each luminance) thus yielded means for each luminance-duration combination based on 50 trials.

Results and discussion.—Mean RTs for the 4 *Ss* combined are plotted in Fig. 1. Each data point is thus based on 200 RTs

For the two highest luminances, duration is unrelated to RT over the range studied. For the 30 ft-L flash, there was a 10-msec. increase in RT when its duration was reduced from 5 to 0.5 msec. For the two lowest luminances, stimulus duration has a far more marked effect on RT, RT being obviously an accelerated function of flash briefness.

Our results show that the CD for moderately intense stimuli (3000 and 300 ft-L) is remarkably brief, being less than 0.5 msec. At 30 and at 3 ft-L, CD lies between 2 and 5 msec. For the weakest target, the CD lies between 10 and 25 msec. The present study shows a small decrease in RT as duration increased from 10 to 20 msec. In the previous study, a smaller decrease occurred between 10 and 25 msec., but there was no further decrease when this stimulus was prolonged beyond 25 msec.

These CDs for RT are far shorter than the 100-msec. value previously reported for absolute threshold (e.g., Baumgardt & Hillmann, 1961) or the minimal value of 30 msec. reported by Graham and Kemp (1938) for the incremental threshold at their highest background luminance. The three dependent variables, RT, RL, and DL, are thus differently related to stimulus duration, with the CD being obviously shortest for RT.

Although luminance differences are confounded with test days (i.e., only one luminance was studied in a given test session), the effect of luminance on RT is pronounced and is apparent at all durations studied. That RT decreases when luminance is increased is consistent with earlier findings (see Woodworth & Schlosberg, 1954). But the form of the relation between luminance and RT will depend on stimulus duration *unless* each stimulus duration is greater than the CD. In other words, our data could be replotted to display six different luminance-RT functions, one for each flash duration.

Our results show that although the overt response to a target flash may not appear until much later, the minimal latency of that response is determined very shortly after stimulus onset. The finding that increasing duration may cease to be effective long before the criterion response appears parallels the classical observation of this fact made by Hartline (1934). The fact that RT is determined by so brief a "package" of luminous energy is consistent with our earlier finding that RT is independent of the growth (with duration) of phenomenal brightness. In addition, it helps to explain why retroactive (metacontrast) masking of a flash does not affect its RT (Fehrer & Raab, 1962).

References

Baumgardt, E., & Hillmann, B. Duration and size as determinants of peripheral retinal response. *J. Opt. Soc. Amer.*, 1961, **51**, 340–344.

Fehrer, E., & Raab, D. Reaction time to stimuli masked by metacontrast. *J. exp. Psychol.*, 1962, **63**, 143–147.

Graham C. H., & Kemp, E. H. Brightness discrimination as a function of the duration of the increment in intensity. *J. gen. Physiol.*, 1938, **21**, 635–650.

Hartline, H. K. Intensity and duration in the excitation of single photoreceptor units. *J. cell. comp. Physiol.*, 1934, **5**, 229–247.

Raab, D., Fehrer, E., & Hershenson, M. Visual reaction time and the Broca-Sulzer phenomenon. *J. exp. Psychol.*, 1961, **61**, 193–199.

Woodworth, R. S., & Schlosberg, H. *Experimental psychology.* (Rev. ed.) New York: Holt, 1954.

(Received June 30, 1961)

Journal of Experimental Psychology
1962, Vol. 64, No. 3, 327–328

SUPPLEMENTARY REPORT: MEANINGFULNESS AS A DIFFERENTIATION VARIABLE IN THE VON RESTORFF EFFECT[1]

HAROLD ROSEN, DONALD H. RICHARDSON, AND ELI SALTZ

Center for the Study of Cognitive Processes, Wayne State University

The present study was designed to test an aspect of the theory for the von Restorff effect proposed by Saltz (1960). This theory defines a ***differentiation*** construct in terms of two variables: (*a*) Similarity between an item and other items in a list, and (*b*) amount of prior reinforcement (e.g., "familiarization") of the item. Isolation techniques typically involve differentiation of an item by reducing its similarity to other items in the list. However, if all the items in a list are already highly differentiated, reduction of similarity would be expected to have a smaller effect than if

[1] The present study is part of a program of research, supported by a National Science Foundation grant to Eli Saltz, concerned with the influence of differentiation on verbal learning.

TABLE 1

MEAN INTRALIST RANK AND MEAN NUMBER OF CORRECT ANTICIPATIONS OF THE ISOLATED AND CONTROL TERMS OVER 15 LEARNING TRIALS

List	Item	Ranks		Correct Anticipations	
		Mean	*SD*	Mean	*SD*
Low *m*	Isolated	4.53	2.24	5.73	3.53
	Control	6.53	1.62	2.12	2.95
High *m*	Isolated	5.00	1.63	7.70	3.94
	Control	5.45	1.89	6.48	3.75

the other items in the list were low in differentiation.

This deduction was tested by comparing the effects of isolation of an item in a serial list of high-meaningfulness items with the effects obtained by isolating an item in a serial list of low-meaningful items, since Noble (1953) has demonstrated that meaningfulness is related to amount of prior experience with an item. Isolation was accomplished by typing one word of the list in red, all the other words being typed in black.

Method.—The *Ss* were 132 students in introductory psychology at Wayne State University, and were randomly assigned to conditions except that *Ns* of conditions were made equal.

Two basic lists were used, each consisting of nine items taken from Noble's (1952) meaningfulness scale. One list consisted of low-meaningful items (*m* range = 1.05 to 1.50): MEARDON, BYSUSS, VOLVAP, LATUK, GOKEM, POLEF, SAGROLE, WELKIN, NARES. The high-meaningful items (*m* range = 7.39 to 9.61) in a second list were: INSECT, JEWEL, HEAVEN, OFFICE, WAGON, DINNER, MONEY, ARMY, KITCHEN. The item in Position 5 was typed in red and served as the isolate in the experimental condition. All other items were typed in black. The control lists were typed entirely in black. Each of the two basic lists was organized into six different random orders, with a different term serving as isolate in each order. An *S* learned one of these orders as a serial list.

Items were exposed on a Lafayette memory drum at a 2-sec. rate with a 4-sec. intertrial interval. After an initial trial in which *S* pronounced aloud the items in order of their appearance, *S* was given 15 anticipation trials.

Results.—The relative effects of isolation for high- and low-meaningful material were evaluated by determining the effect of isolation on each *S*'s serial position curve. For each *S*, the items in the list were ranked from greatest number of correct anticipations in 15 trials (Rank 1) to least correct anticipations (Rank 8). The mean ranks of the isolated and control items, summarized in Table 1, indicate that isolation has a much greater relative effect in a list of low-*m* than a list of high-*m* items. The main effects due to isolation are significant beyond the .001 level ($F = 14.36$, $df = 1/128$) and the Isolation × Meaningfulness interaction is significant beyond the .025 level ($F = 5.77$, $df = 1/128$). Table 1 shows that the interaction effect also occurs for mean number of correct anticipations over 15 trials. However, the anticipation data are less persuasive than the rank data, since the anticipation data could be an artifact due to the rapid learning of high-*m* items, both isolated and control. In brief the original prediction based on the differentiation formulation is sustained.

Saltz and Newman (1959) found that isolation resulted in an increased tendency for the isolated term to be emitted as an intrusion response to other stimuli in the list. The intrusion data in the present study were analyzed by means of a chi square test since the data were markedly skewed: the median number of intrusions was zero in the low-*m* control groups, and close to zero in the other groups. For low-*m* lists, 58% of the *Ss* in the isolated condition emitted the isolated term as an intrusion at least once during learning; only 33% of the control *Ss* emitted the control item. This difference is significant ($P < .05$). The corresponding percentages in the high-*m* conditions are 58% and 67%, producing a nonsignificant difference opposite in direction from that for low *m*. Saltz (1960) has hypothesized that differentiation increases the tendency for a response to be emitted, and that this tendency is basic to the relatively high intrusion rate for the isolated term. In terms of this position, the fact that the high-*m* isolation condition produces no greater tendency toward intrusions than the high-*m* control condition is consistent with the position stated previously in this paper: in a list of highly meaningful items, the terms are already relatively differentiated, and so isolation will contribute relatively little additional differentiation.

REFERENCES

NOBLE, C. E. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421–430.

NOBLE, C. E. The meaning-familiarity relationship. *Psychol. Rev.*, 1953, **60**, 89–98.

SALTZ, E. Similarity and differentiation in verbal learning. Paper read at Psychonomics meetings, Chicago, 1960.

SALTZ, E., & NEWMAN, S. E. The von Restorff isolation effect: Test of the intralist association assumption. *J. exp. Psychol.*, 1959, **58**, 445–451.

(Received August 1, 1961)

CONTENTS—*Continued*

The Course of Emotionality in the Development of Avoidance: H. S. Hoffman and M. Fleshler 288

On the Relations among Some Factors That Contribute to Estimates of Verticality: C. R. Curran and H. L. Lane 295

The Perspective Illusion: Perceived Size and Distance in Fields Varying in Suggested Depth, in Children and Adults: J. F. Wohlwill 300

Stimulus Generalization as a Function of UCS Intensity in Eyelid Conditioning: J. J. Porter 311

Experimental Extinction as a Function of Number of Reinforcements: J. R. Ison 314

Eye-Movement Latency, Duration, and Response Time as a Function of Angular Displacement: A. E. Bartz 318

Supplementary Report: Direction of Change in CS in Eyelid Conditioning: F. A. Logan and A. R. Wagner 325

Supplementary Report: The Effect of Stimulus Duration and Luminance on Visual Reaction Time: D. Raab and E. Fehrer 326

Supplementary Report: Meaningfulness as a Differentiation Variable in the von Restorff Effect: H. Rosen, D. H. Richardson, and E. Saltz 327